PHYSICAL CHEMISTRY AND BIOPHYSICS

FOR

STUDENTS OF BIOLOGY AND MEDICINE

BY

MATTHEW STEEL, Ph.D.

Professor of Biological Chemistry
The Long Island College Hospital
Brooklyn, N. Y.

PREFACE

In preparing this textbook the aim has been to restrict its contents to such a degree as to make it suitable for a course continuing through at least a third of an academic year. Owing to the limited time at his disposal, it is quite impossible for the average student of medicine to find out for himself what is for him important and essential amongst the immense mass of physico-chemical data at his disposal. The necessity therefore has arisen for treating the rudiments of physical chemistry in a manner especially adapted to the requirements of medical and biological students, and it is with this object that the author has written the present text. The treatment of the subject is based on the course in physical chemistry pursued by medical students at the Long Island College Hospital, and special attention is given to the physiological and medical bearings of the subject.

It is the belief of the author that, in order that a student may get the most out of a technical science like physical chemistry, his interest must be aroused at the outset. He must be made to perceive that the subject in question has a direct bearing on the problems in which he is interested. Accordingly, a brief introductory chapter is given, which serves to indicate the importance of physics and physical chemistry in the study of vital phenomena.

Considerable attention is given to the modern conception of the nature and structure of matter because the composition of the atom is the keynote of the properties of matter in all the various states of aggregation. Without a knowledge of the present-day theories in regard to the nature and structure of the atom, it is impossible for the student to comprehend such topics as the cause of ionization, the nature of amphoteric electrolytes and their iso-electric points, hydrogen-ion concentration, Donnan's membrane equilibria theory, oxidation-reduction phenomena in the absence of free oxygen or hydrogen, etc.

In the chapter on the Colloidal State of Matter the researches of

Jacques Loeb on proteins have received special attention owing to the far-reaching influence that Loeb's discoveries and conclusions have on all biological and medical problems.

The treatment of the subject throughout the course is intended to lead the student to feel that physical chemistry will enable him to possess a more lucid perception of the chemical changes that occur in living matter in health and in disease. To this end, the author, when possible, has illustrated the application of physico-chemical principles by examples taken from the fields of biology, physiology, and medicine.

MATTHEW STEEL

BROOKLYN, N. Y.
Nov. 23, 1927

CONTENTS

PHYSICAL CHEMISTRY AND BIOPHYSICS FOR STUDENTS OF BIOLOGY AND MEDICINE

CHAPTER I

GENERAL INTRODUCTION

If one attempts to study living matter from a chemical standpoint only, one soon encounters almost insurmountable difficulties. For, directly one begins to analyze living matter by the standard methods that are applicable to minerals and soils, etc., it ceases to be living matter. The reagents which are employed for such purposes immediately " kill " it, that is to say, effect a suspension of some or all of the properties which characterize living matter. (Movement, growth, reproduction, respiration and irritability are the properties of living matter, or protoplasm.)

If one boils living matter or treats it with strong acids or alkalies, etc., one will probably be able to separate many of the chemical constituents from it and to identify them, but the products which the analysis yields bear only a remote relationship to those which were originally present in the living organism.

On the other hand, physico-chemical methods allow one in many cases to follow quantitatively the influence of temperature and foreign substances upon these products without effecting their destruction. In other words, in order to find out what is actually occurring in living matter one must employ methods of investigation that will not involve the cessation of the very processes one desires to investigate. One must use *static* and not *dynamic* methods of mensuration in studying living matter. Of course it is necessary to use dynamic methods to obtain knowledge of the decomposition products of living matter.

It is evident, therefore, that the necessary data for the study and interpretation of vital phenomena cannot be obtained by the ordinary

chemical methods only, i.e., that while these methods are absolutely essential they must be supplemented by the laws of physics and by physico-chemical methods. The most casual knowledge of the subject will convince one of the value of such laws as, the *law of mass action*, which permits of an insight into the course, velocity and equilibrium of chemical reactions; the *theory of osmotic pressure*, which has revealed to us the common properties possessed by any series of solutions independently of the nature of the dissolved substances; and the *theory of electrolytic dissociation of certain substances;* the foundation of which resides in the partial dissociation of molecules into electrically charged atoms or groups of atoms acting *en masse*, called ions.

The most striking tendency of modern physiology is the adoption of chemical and physical methods of analysis. In fact, so great is the volume of work that comes under the headings of chemical physiology and physical physiology, that new teaching and research departments have been formed in these branches of the mother science. These departments are now called *biochemistry* and *biophysics*, respectively.

Biochemistry may be defined as the study of the chemical and physico-chemical processes that play a part in the life phenomena of plants and animals. In the past biochemistry has dealt with comparatively simple problems, mainly with the analysis of the materials of which the cell is composed after it has been killed—in a word, with chemical anatomy. The in-goings and out-goings of the body have been analyzed. In general it may be said that the progress of biochemistry has been along the same path that has been followed by other branches of experimental science. First a descriptive phase and then a quantitative phase is the general procedure. Biochemistry has followed this path, and in doing so has developed methods of a much more delicate type than are usually called for in other branches of chemistry. The biochemist often has to deal with the merest traces of materials. For instance, in working with the accessory food factors—the so-called " vitamins "—the quantities are so infinitesimal as to demand the finest of experimental chemical and physical skill.

Chemistry and physics have made great strides in the past two decades and these advances in knowledge and in methods have been reflected in the biological sciences. Modern physics is getting inside the structure of the atom; chemistry, inside the structure of the molecule; and biochemistry is getting inside the structure of the living cell by employing physico-chemical procedures.

The study of the structure of the atom by the physicists has revealed the secret of oxidation and reduction in general. The atom has been shown to be made up of varying quantities of positive electricity, called "protons," and of negative electricity called "electrons"; while oxidation has been shown to consist fundamentally in the loss of valence electrons by some atoms or ions, and reduction, the reciprocal gain of such electrons. (These facts will be elucidated in the body of the text.)

The science of biophysics has developed in the same manner as biochemistry. It differs from its sister subject—biochemistry—in technique, but not in purposes. Many of the characteristic reactions of living matter may be explained in terms of pure physics. For example, hydrostatics and hydrodynamics illuminate the technical procedures employed in the measurement of blood pressure and in the interpretation of sphygmographic records; the intelligent use of certain instruments of precision, such as the laryngoscope and ophthalmoscope, can only be obtained by an extensive study of optics. Likewise, the use of the cathode-ray oscillograph for recording the action current of nerves, the development of calorimetric and electric means of studying the energy exchanges of animals, the use of electrometric methods of measuring ionic concentrations, the investigation of membrane equilibria and surface forces—all of these require a skillful adaptation of physical technique to the special biological problems. The actions of heat, light, and friction on our sense organs can only be solved by the use of the methods and instruments of physics. The energy transformations in living matter can all be explained by applying the laws of thermodynamics.

Medicine has never denied her dependence upon the advances in the basic sciences, and so it is not surprising to find that with new ideas in chemistry or physics there have come corresponding discoveries in medicine. But there have been many disappointments, owing to the fact that attempts have been made to solve certain medical problems by physical or chemical principles and methods that were not suitable to the problem under consideration. Often physical methods alone, or chemical methods alone, were inadequate in solving the intricate problems of physiology or pathology, but when physical and chemical methods and principles were combined the goal was quite frequently attained. It is not strange, then, to find that the great advances in medicine have taken place since the amalgamation

of the sciences—physics and chemistry—into the new "hybrid" science—physical chemistry.

This amalgamation was undoubtedly brought about through researches in physics in the field of thermodynamics. As Pauli has pointed out: "The transformations in energy in chemical reactions have two general sources. As is well known, every change in the state of aggregation is accompanied by either an absorption or an evolution of heat. Since changes in physical state often accompany a chemical reaction, these constitute, therefore, one of the sources of the transformations in energy accompanying this reaction. A second is found in the chemical reaction itself. The synthesis or analysis of a substance is accompanied by a thermal change which may have either a positive or negative value. To illustrate this we cite the formation of a salt from an acid and a base with the development of the so-called heat neutralization; or the decomposition of a salt into its components with a using up of electrical energy. All these metamorphoses in energy constituted a fruitful field of work, in which medicine also took a part."

One of the most striking aspects of the development of modern physiology and modern medicine is the part played by considerations derived from physical chemistry. The osmotic pressure of solutions; the action of various inorganic ions arising from the electrolytic dissociation of acids, bases, and salts, especially the hydrogen ion; catalytic phenomena; and the maintenance of the neutrality necessary to the normal function of the cell, together with the properties of matter in the colloidal state, are especially worthy of mention. Again, to quote Pauli: "All living matter is made up of colloidal and crystalloidal material, and there exists no life process that is not accompanied by changes in the colloidal and crystalloidal substances. And the physico-chemical laws which govern the crystalloids and colloids reappear in the numerous properties of living matter."

One of the fundamentally important steps in the physico-chemical characterization of living matter was taken when Graham divided all bodies into colloids and crystalloids, according to their physical characteristics. From this division there has arisen the conception that all living matter is of necessity connected with the existence of a colloidal ground-substance in which all changes take place.

About 1860, Thomas Graham, while studying the general problem of diffusion, recognized that chemical substances differ greatly in the

rate at which they diffuse through solvents of various kinds. On the basis of this observation he made a distinction between those which diffuse slowly or not at all and those which diffuse very rapidly. As the former are for the most part amorphous, and since ordinary glue is an example of this class, he called them *colloids* (the Greek word for glue is *kolla*). The group that diffused readily he called crystalloids, for such easily crystallizable substances as cane sugar, common salt, and urea are found in this group.

Since the time of Graham it has been found that most substances can be obtained in either the colloidal or the crystalloidal state, the state of aggregation depending on the manner of manufacture.

It is likewise now common knowledge that most animal and vegetable fluids and solids are primarily colloidal in character. Or, to be more exact, the greater part of living matter exists in the colloidal state intermixed with a small quantity of crystalloids. It is for this reason that an exhaustive study of colloidal chemistry is of vital importance for the student of medicine. Two or three specific illustrations of the nature of the problems one has to face must suffice here. All colloidal systems, and particularly those of the body, are extremely sensitive to chemical and physical changes in their environment. Slight changes in the amount of acid in a given system or slight changes in the temperature will often cause the colloids to coagulate, and therefore to lose their normal properties. The blood is peculiarly sensitive to acidity. Very slight changes in the reaction of the blood will result in profound metabolic disturbances; it is not surprising, therefore, to find that nature has supplied a method of keeping this fluid almost constant in reaction. Although an animal may be constantly consuming large quantities of acids and alkalies in its food, the blood remains practically neutral in reaction. This is especially true in healthy animals. The neutrality of blood was not explainable until physical chemistry had advanced sufficiently to give us the theory of " buffers," a class of substances that are capable of reacting with either acids or alkalies, and in such a manner that the fluid containing them will undergo but slight change in reaction with the addition of new portions of acids or alkalies. Now, the theory of buffers could not have been conceived before we had the theory of electrolytic dissociation—a great field that was unknown before the year 1887. Likewise, the interesting problem of selective absorption by living colloidal membranes and organs is only explainable by principles that have been

developed in the field of physical chemistry. We shall see later on in the text that certain living cells are permeable to some of the ions of the salts surrounding them, but not to others. The same cells do not possess this selective ability after death.

The protective action of certain colloids, the formation of gels, swelling, surface-tension phenomena, and adsorption are other physico-chemical fields of investigation that have had a great bearing on the solution of numerous physiological and pathological processes. For example, the formation of cell membranes is largely explained on the basis of adsorption, and the same is true of the neutralization of toxins by anti-toxins.

CHAPTER II

THE NATURE AND STRUCTURE OF MATTER

1. The Development of the Atomic Theory.—Some five hundred years before the Christian era, two philosophers independently gave birth to the atomic theory of matter. One was a Hindu and the other was a Greek. The Hindu philosopher, Kanada, advanced the idea that matter is made up of very small indivisible particles which are in a constant state of motion, moving in every direction. Almost the same ideas were held by the Greek philosopher, Leucippus, who is known as the founder of the Greek atomistic school. The atomic theory was further developed by Democritus, a pupil of Leucippus. About the same time, Empedocles, another Greek philosopher, introduced the notion that everything consists of four " elements "—earth, water, air, and fire. He also, on purely philosophical grounds, originated the idea of the conservation of matter. He stated that, " nothing can be made out of nothing, and it is impossible to annihilate anything. All that happens in the world depends on a change of form and upon the mixture or the separation of bodies. An everlasting circulation is characteristic of nature." Plato and Aristotle had somewhat similar ideas; they asserted that the " elements " might be converted into one another. (There is no reason to believe that the term " element " was used by the ancient philosophers in the same sense as it is to-day.)

While the atomic theory was rather generally accepted by philosophers in different parts of the globe, owing to the paucity of experimental data it made little progress for more than two thousand years after the death of Empedocles.

About the middle of the seventeenth century the theory was revived by Gassendi, a prominent physicist of that time. It was also about this time that it was accepted by such eminent scientists as Robert Boyle and Sir Isaac Newton. We are indebted to Boyle for the modern conception of the term " element."

7

The atomic theory was also stated in extraordinarily modern terms in 1743 by Lomonosoff, a Russian physical chemist; however, owing to the fact that very few foreigners speak or read the Russian language, his work has only recently been " discovered " (1904).

It is to John Dalton, an English scientist, that the world owes its present conception of the atomic nature of matter. Dalton, in 1801, used the atomic theory to explain the diffusion of gases, and in 1803 he enunciated his theory of the structure of matter and of chemical combinations. In 1808 he offered more convincing data. Dalton may be said to have quickened the dead dogma of former times into a living hypothesis. It was he who first applied the idea of discrete particles quantitatively to chemistry. He contended that all the atoms of an individual element must be of the same size, weight, and shape. Otherwise it would be possible to obtain different types of the same element having different properties. By his quantitative experiments he was able to show that the combining proportions are exactly those which would exist if each element consisted of a great number of particles, all the particles of any one element being exactly alike, and that the atoms of different elements have different properties—weight, shape, size, affinity, etc. He estimated the relative weights of the atoms of quite a few elements. He showed that compounds are formed by the union of atoms of different elements in simple proportions— 1 : 1; 1 : 2; 2 : 1; 2 : 3, etc.

While Dalton's atomic hypothesis, especially his conception of atomic weights, is not exactly that which prevails in modern chemistry, it was such a clear-cut and practical presentation of chemical combination that it was nearly universally adopted, and it pointed the way towards new discoveries.

About the same time that Dalton was working on the combining weights of atoms, Gay-Lussac announced his theory of the combining volumes of gases. He observed that gases entered into chemical combinations in simple proportions by volume, provided the volumes of the different gases were measured under the same external conditions of temperature and pressure. Now, Dalton had come to the conclusion that gases (in common with other substances) enter into chemical combination in simple atomic proportion. There must, therefore, be some simple relationship between the atoms of gases and the spaces they occupy. Dalton had also conceived the idea that equal volumes of different gases contained the same number of atoms. In other words,

by combining the two ideas, it became evident that the weights of equal volumes of different gases were proportional to the weights of their atoms. Many difficulties were encountered when attempts were made to correlate the theories with the observed facts. The problem of determining atomic weights was undertaken by Berzelius, he having become convinced of the importance of the connection between Dalton's theory and Gay-Lussac's law of combining volumes. He came a little nearer to overcoming the difficulty than Dalton. His researches on water and hydrochloric acid will illustrate the point.

Dalton's supposition that an *atom* of water consisted of one atom of hydrogen and one atom of oxygen was at variance with the belief that equal volumes of gases consisted of equal numbers of atoms, for experiments showed that the combining volumes of hydrogen and oxygen were in the proportion of 2 : 1. Berzelius, therefore, asserted that an atom of water is composed of two atoms of hydrogen and one of oxygen. When he experimented with hydrochloric acid, however, a difficulty arose which he failed to overcome. In this case, one volume of hydrogen and one of chlorine combine to form two volumes of hydrochloric acid. This would imply that one atom of hydrogen combines with one atom of chlorine to form one " atom " of hydrochloric acid. These facts were incompatible with the atomic theory as then understood. There were not enough atoms of hydrogen to go around. Consequently, it must be wrong to assume that equal volumes of gases contain the same number of atoms. Avogadro, in 1811, solved the difficulty. He distinguished between the elementary atoms and the small particles of a gas. Each gas particle was assumed to be an aggregate of a definite number of atoms. Avogadro called the groups of atoms " molecules." On the assumption that each molecule of an elementary gas contains the same number and kind of atoms, and that a molecule is the smallest particle of an element or compound that exists free in a gas, he then introduced his modification of Gay-Lussac's hypothesis, namely, " Equal volumes of all gases, at the same temperature and pressure, contain the same number of molecules." Up to the present time, no one has been able to disprove this hypothesis. By assuming that the molecules of oxygen, hydrogen, and chlorine gases are made up of two atoms and that hydrochloric acid molecules are made up of one atom of hydrogen and one of chlorine, also that a molecule of water is made up of two atoms of hydrogen and one atom of oxygen, most of the difficulties of Dalton, Gay-Lussac and Berzelius

were overcome. However, nearly fifty years passed before Avogadro's ideas were generally accepted. During the second half of the nineteenth century the atomic theory was greatly advanced by the work of a number of scientists on the " kinetic theory of gases." It was shown that many of the properties of gases can be explained on dynamic principles if the gas is considered as a multitude of minute particles in rapid motion.

During the last few decades a number of new discoveries have been made, new instruments have been invented, and methods of investigation have been so refined that a powerful light has been thrown on the subject. The study of electrified atoms has been especially fruitful. In short, the data that have been obtained in the past few years, from so many chemists and physicists working on different phases of the subject, have been in such perfect accord that the atomic and molecular structure of matter is now accepted as an established fact.

2. Atoms and Molecules.—According to the modern theory of matter, all elementary substances are complex structures made up of minute particles called atoms, which are all alike. Each atom, as we shall see later, is made up of a definite number of smaller particles. An atom, however, cannot be broken up into its constituents by *chemical processes*. Therefore, there must be as many kinds of atoms as there are elementary substances.

As was indicated in Art. 1, a careful study of the chemical and physical properties of matter led Avogadro to the conclusion that the reactive particles of matter are not the atoms, but rather groups of two or more atoms, which he called molecules. This idea still holds.

If, on investigation, a given type of matter is found to be composed of molecules that are all alike, it is called a *pure substance*. On the other hand, if a given material is composed of more than one kind of molecules—it is called a *mixture*.

3.—Simple Substances, or Elements, and Compound Substances. When the atoms making up the molecules of a pure substance are all alike, the substance is called an *element*. A compound is a pure substance the molecules of which are composed of more than one kind of atoms. Boyle, and later Lavoisier, defined an element as a substance that has not so far been split into anything simpler. Lavoisier's conception of an element was so clear-cut that one feels justified in quoting his words. " If we apply the term element or principle to bodies to express our idea of the last point which analysis is capable

of reaching, we must admit, as elements, all substances into which we are able to reduce bodies by decomposition. Not that we are entitled to affirm that these substances which we consider simple, may not themselves be compounded of two, or even a greater number of simple principles; but since these principles cannot be separated, or rather, since we have not hitherto discovered the means of separating them, we must regard them as simple substances, and we ought never to suppose them compounded until experiment and observation have proved them to be so." In the course of time it came to be recognized that many substances that Lavoisier considered to be elements were really compounds. For example, lime, baryta, magnesia, and alumina were supposed by him to be elements.

There are some eighty-seven substances that have so far resisted further decomposition by *chemical processes*. These elementary substances have certain well-defined properties which are not possessed by compounds. These distinctive properties can be obtained from textbooks on general chemistry.

Recent investigations have shown that some of the elements, namely, the radioactive elements, are continuously undergoing spontaneous disintegration, with the final production of other elements. However, this spontaneous disintegration must not be confounded with *chemical action*, which is the name given to the process whereby molecules are either formed or destroyed.

4. The Arrangement of the Atoms in the Molecules: Allotropism and Isomerism.—There are several elementary substances and many compounds which are composed of the same atoms and the same relative quantities of these, but which have different molecular weights and different properties. For example, ordinary oxygen is known to be composed of molecules which contain 2 atoms of oxygen. There is also another substance, known as ozone, which is made up of molecules which contain 3 atoms of oxygen. Each of these substances will, on combining chemically with some other element, yield the same end product. Another well-known example of this phenomenon is that of sulphur. There are several substances having distinctive properties, each of which on uniting with oxygen will yield 2 grams of sulphur dioxide for each gram of substance taken. This proves that each of the different forms of sulphur is but a modification of one element. It is likewise probable that yellow and red phosphorus are composed of molecules containing different numbers of atoms of the same element.

There are other cases known in which the molecules are composed of the same number of atoms, all of the same kind, yet the substances have different physical properties. Again, sulphur serves as a good example. Di-atomic, tetra-atomic, hexa-atomic, and octa-atomic sulphur are known, and there are also two distinct octa-atomic forms. One of the types crystallizes in rhombic crystals, and the other in monoclinic crystals. They have different freezing-points and boiling-points. The following structural formulæ have been suggested, but not proved:

$$S\equiv S-S\equiv S$$
$$|\qquad\qquad|$$
$$S\equiv S-S\equiv S$$

Rhombic sulphur or
D-sulphur

$$S=S=S=S$$
$$\|\qquad\qquad\|$$
$$S=S=S=S$$

Monoclinic sulphur or
B-sulphur
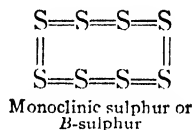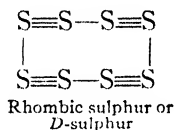
That property in virtue of which the same element exists in more than one form, each of the different forms having distinctive properties, is called allotropy, from the Greek (*allos*) another, and (*tropos*) shape. The less common form is often called an " allotropic " modification of the other, and the phenomenon is known as *allotropism*.

Just as in the case of elementary substances, but more often, we find compound substances having the same atomic composition occurring in different structures. Such compounds are called *isomeric substances*, or *isomers*. The general phenomenon is known as *isomerism*.

Isomeric substances often have different properties. This is especially true among the hydrocarbon derivatives. For example, at least two substances are known that have 2 atoms of carbon, 6 atoms of hydrogen, and 1 atom of oxygen (C_2H_6O). One of these compounds is ordinary alcohol, which, as is well known, is a liquid that boils at 79° C. The other is methyl ether, which is a gas at ordinary temperatures. It boils at about −24° C. By studying the manner in which alcohol can be built up from its elements and by observing its properties, we are able to assign to it the structural formula:

$$\begin{array}{c} \text{H}\quad\text{H} \\ \text{H—C—C—OH.} \\ \text{H}\quad\text{H} \end{array}$$

The other substance is found by similar processes to have the following structural formula:

$$\begin{array}{c} \text{H}\qquad\text{H} \\ \text{H—C—O—C—H.} \\ \text{H}\qquad\text{H} \end{array}$$

In fact, the study of " organic chemistry " is largely a study of isomeric compounds. For instance, the organic chemist is familiar

with 157 different compounds which are known to have the "empirical" formula: $C_{10}H_{16}$.

The inorganic chemist is also acquainted with many isomeric compounds. Thus, there are two distinct compounds containing 4 atoms of hydrogen, 2 of nitrogen, and 3 of oxygen. The general properties of these two substances are so very different, that it is only by assuming them to have a different arrangement of the atoms in the molecule that one is able to offer a plausible explanation of the differences that are found in their properties. One of the compounds is called ammonium nitrate, and the other hydroxylamine nitrite. The probable structural formulæ of the two compounds are:

$$ \begin{matrix} H \\ H \\ H \\ H \end{matrix} \!\!\! > N - O - N \!\!\! \begin{matrix} O \\ O \end{matrix} \qquad\qquad \begin{matrix} H \\ H \\ H \\ HO \end{matrix} \!\!\! > N - O - N = O $$
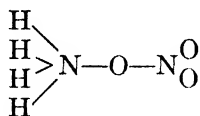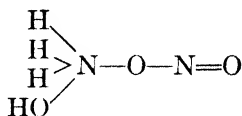
<div align="center">Ammonium nitrate   Hydroxylamine nitrite</div>

There is another class of isomeric compounds which have the same composition and are alike in every other property with the exception of that which they exhibit on the rotation of the plane of polarization of light. Such compounds are usually derivatives of carbon and possess at least 1 carbon atom whose 4 valences are satisfied by 4 different radicals (or atoms). This connection between optical activity and atomic structure was first discovered by Le Bel, but before he published his discovery it was independently made by van't Hoff. Van't Hoff named the central carbon atom of the group an asymmetric carbon atom.

Since the event of the noteworthy researches of Le Bel and van't Hoff in 1874, the phenomenon is known as " space " or " stereoisomerism."

A study of lactic acid will serve to illustrate the more important features of stereoisomerism. Lactic acid (1-hydroxy propionic acid) is known in three forms. One form, dextrolactic acid, rotates the plane of polarized light to the right; the second form, lævolactic acid, rotates the plane of polarized light to the left; while the third form, ordinary lactic acid, is optically inactive. The third form has been shown to consist of equal quantities of dextrolactic acid and lævolactic acid molecules; in this case the two constituent acids neutralize each other in their action on polarized light. The following struc-

tural formula will indicate which carbon atom is the asymmetric one:

$$CH_3$$
$$|$$
$$H-C-OH$$
$$|$$
$$COOH$$
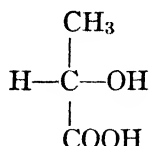
It will be noted that one of the carbon atoms has each of its 4 valencies satisfied by 4 different radicals. If the asymmetric carbon atom is represented as lying in the center of a tetrahedron, of which the apices represent the positions of the 4 radicals attached to the asymmetric carbon atom, it becomes easy to understand the different arrangement of the radicals causing the asymmetry. This is illustrated in the following figure:
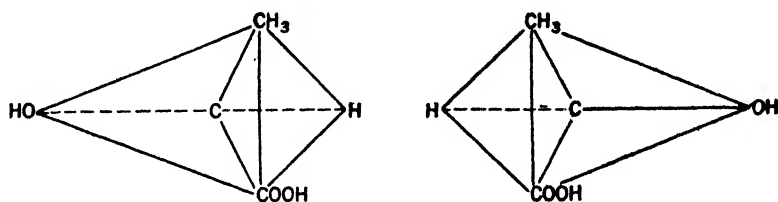
FIG. 1.

If we now try the effect of interchanging these groups in all possible ways, it will be found that two and only two different arrangements are possible. Further, the two tetrahedrons are not superimposable, but if we hold the tetrahedron representing one combination before a mirror, the image in the mirror will be a facsimile of the other possible arrangement.

The two isomers shown in Fig. 1 obviously have the same constitution, but differ in the spatial arrangement of the groups in the molecule, i.e., they have a different " configuration."

Several other varieties of isomerism could be cited, but enough has been said to illustrate the principle that many substances are essentially different because the arrangement of the atoms in their molecules are different.

5. The Atom a Complicated Structure.—We have seen that the structure of a molecule may be very complex, especially if it happens to be composed of a large number of different atoms. Moreover,

millions of chemical reactions have confirmed Dalton's conception of the atom, namely, that an atom is the smallest particle of an element that can enter into or be expelled from chemical combination. (The word atom comes from the Greek words *a*, meaning " not," and *temo*, meaning " cut,"—" that which cannot be divided.") Yet, paradoxical as it may seem, modern research has demonstrated that the structure of the atom is, perhaps, more complex than that of the molecule. In brief, the study of the discharge of electricity through rarefied gases and the study of the radioactive substances have not only supplied strong experimental support of the fundamental ideas of the atomic theory of matter, but they have also indicated that the atom itself is not the smallest unit of matter, as was formerly supposed, but is a complicated structure made up of a number of smaller bodies. Every atom has been shown to contain a considerable number of very small distinct particles, called *electrons* or corpuscles. These particles have been shown to be nothing but free charges of negative electricity. They are all exactly alike, no matter what atom they may come from, and each constitutes the smallest quantity of electricity known to exist. Atoms, therefore, are partly made up of corpuscles of negative electricity. Of course, to prevent the spontaneous disintegration of the atom by the mutual repulsion of other similarly charged particles, it is necessary to assume the existence of an equal amount of positively charged electricity. In other words, it is assumed that each atom of any element consists of a large number of electrons, all negatively charged, which are held together by positive electricity equivalent in amount to the sum of the negative charges of all the electrons present, so as to form an electrically neutral atom. This theory of the composition of matter, which we owe largely to Sir J. J. Thomson, has enabled the chemist to explain the periodic law, valency, chemical affinity, chemical action, etc., in terms of known facts.

Before the atomic system is described in detail, the foregoing statement must be proved. This can, perhaps, be done most satisfactorily by briefly reviewing the more important experimental discoveries that have led up to this conception.

6. Birth of the Electron Theory.—It is generally conceded that the modern conception of the nature of matter had its birth nearly a .century ago in the work of Michael Faraday upon electromagnetism and the conduction of electricity in solution. His discovery was of such transcendent importance that it was bound sooner or later to

lead up to our present ideas upon matter and electricity. It was that whenever two metals or other elements of the same valency are deposited or evolved in the same electrolytic cell, the amounts of electricity consumed, as measured by Lane's unit jar or other instruments, are inversely proportional to the atomic weights of the elements. Or, in other words, that the electricity attached to every atom of a given valency is the same; that is to say, if a metal is divalent its atom is associated with twice the atomic quantity of electricity of a monovalent atom.

Commenting upon this discovery in his Faraday lecture, Helmhotz said: " If we accept the hypothesis that elementary substances are composed of atoms, we cannot avoid the conclusion that electricity, positive as well as negative, is divided into definite elementary portions which behave like atoms of electricity."

A number of years later, an Irish physicist, G. J. Stoney, at the Belfast meeting of the British Association, drew attention to the " atom of electricity " as one of the three fundamental physical units of nature (the others being the velocity of light and the constant of gravitation), and gave an approximate calculation of its value. He said: " Finally nature presents us, in the phenomena of electrolysis, with a single definite quantity of electricity, which is independent of the particular bodies acted on. To make this clear I shall express Faraday's Law in the following terms: which as I shall show, will give it precision, viz., for each chemical bond which is ruptured within an electrolyte a certain quantity of electricity traverses the electrolyte which is the same in all cases."

He called this quantity of electricity an *electron*.

He calculated the actual charge by dividing the quantity of electricity required for the electrolysis of 1 c.c. of hydrogen by the number of hydrogen atoms in 1 c.c. as given by Lochschmidt, and found 10–20 " amperes " (now called absolute electromagnetic units of quantity). This figure compares well with the latest value for the electron.

7. Cathode or Lenard Rays.—A great impetus to the solution of this problem was initially given by the experiments on the *cathode rays*.

W. Hittorf, in 1869, showed that if a body—say, a Maltese cross made of mica—is place between the cathode and the anode in a high-vacuum tube, i.e., a tube containing a highly rarefied gas, a true

shadow appears on the glass; the shape of the cross shows that something must travel from the neighborhood of the cathode in straight lines. This " something," which causes the phosphorescence of glass, was called by E. Goldstein (1876) *cathode rays*. Hence, the cathode rays travel in straight lines normal to the surface of the cathode; and they will cast a well-defined shadow if a solid object be placed between the cathode and the wall of the vacuum tube.

The cathode rays were studied in detail by Sir William Crookes. He showed, in 1879, that the cathode rays can exert mechanical pressure. He did this by arranging the stream of cathode rays so that they would strike the upper vanes of a little paddle wheel which then rolled horizontally along a pair of parallel glass rails, away from the cathode. When the electric current was reversed, the wheel would stop and then revolve in the opposite direction. By directing the cathode rays on different minerals, a beautiful phosphorescent effect was obtained. Again, Crookes showed that the cathode rays raise the temperature of bodies on which they fall. This he did by focusing the cathode stream on platinum by means of a cathode shaped like a concave mirror. The metal became white hot, the heat being sufficient to char diamonds. Crookes also showed that if the cathode stream was allowed to impinge on white rock-salt or lithium chloride these salts assumed a violet color. Hence, cathode rays can produce physical or chemical changes. About 1886, Crookes became convinced that cathode rays are negatively electrified projectiles endowed with a high speed of motion, but he was unable to give positive proof of his ideas.

Hertz discovered that cathode rays can pass through thin metal plates several microns thick, and Lenard, in 1894, showed that they can escape from the tube wherein the discharge takes place through a thin metallic plate strong enough to support the pressure of the atmosphere.

Jean Perrin in 1895 demonstrated that the " cathode rays carry negative electricity with them into a completely closed box and that they are, moreover, deviated in an electric field." He also found that it was also absolutely impossible to separate the electricity from the rays, even by causing them to pass through a thin metallic plate.

About this time the study of the cathode rays was taken up by Sir J. J. Thomson, and in 1897 he, as a result of epoch-making experiments, astounded the world by stating that what Crookes called " radiant matter," i.e., the cathode rays, is a stream of negatively

charged particles or corpuscles which have been formed by *the disintegration of atoms* of the gas in the vacuum tube.

As previously stated, the term *electron* was applied by G. J. Stoney to designate the unit or atomic charge of electricity, and it is this term that is now almost universally applied to the sub-atomic particles supposed to stream from the negative electrode when a discharge is passing through an attenuated gas. No difference can be detected in the corpuscles from the different gases, and hence it is inferred that *these corpuscles or electrons are common constituents of all gases.* If the stream of electrons is directed into an atmosphere of moist air, each electron serves as a nucleus about which moisture collects, and the size of the drops of water formed can be determined by the rate at which the fog which they compose settles under the pull of gravity. Now, since it is easy to measure the total amount of water that condenses, the number of droplets in the fog can be calculated from a knowledge of their size; and, since each droplet corresponds to a single electrical particle, the number of droplets gives a measure of the number of particles that are present, and hence permits the calculation of the charge which they individually bear.

Since we know the amount of electricity carried by each particle in the cathode rays, it is possible to separate the effect of the charge from that of the mass, and hence to ascertain the magnitude of the latter. The most refined measurements of this sort show that the cathode-ray particles, which are now called *electrons*, have a mass which is about one eighteen-hundredth $(\frac{1}{1800})$ that of the lightest known atom, that of hydrogen.

8. Roentgen Rays.—A discovery that has had a direct bearing upon our knowledge of the nature of matter was made by Roentgen, a German physicist, in 1895. He found that when an electric discharge is passed through a Crookes or Lenard tube, there is given off from the tube a hitherto unknown radiation which is characterized by its great penetrability. This radiation will pass through substances that are entirely opaque to ultra-violet rays. These Roentgen rays, or *X-rays*, as they are called, like the cathode rays, excite fluorescence of various kinds, and they affect photographic plates. It is owing to the fact that some kinds of matter, e.g., bones and metals, are less permeable to them, that shadow pictures of such objects can be obtained. They differ from the cathode rays in that they carry no electrical charge and hence they are not deflected either by electrified

bodies or by magnets. Another remarkable property is their ability to ionize gases through which they pass. That is to say, the passing of the X-ray through a gas causes the expulsion of one or more electrons from an atom, leaving the atom with an excess positive charge, since, as was pointed out above, an atom is made up of negatively charged particles (electrons) and a positive nucleus which *just* balances the negative charges present. These atoms that have lost one or more electrons are called " ions," and they are able to take up electrons from any available source. Hence, they can discharge an electroscope.

What the nature of these X-rays may be has been the subject of much study and discussion. Sir George Stokes reasoned that they are a series of pulses in either, emitted at regular intervals, since they are produced by cathode rays striking against the walls of an exhausted tube. The cathode rays, which are highly charged particles, rain down upon the walls of the vacuum tube, and each particle sends up a pulse in the ether. The penetrating power of the rays is explained by assuming that the pulse is lost before there is any harmonious vibration between the ether and the molecules.

The researches of Laue, Bragg, and others, in the past few years, have shown that Stokes' conception of the nature of the X-rays is not quite correct. If it could be proved that the X-rays are a series of regular vibrations in the ether, having extremely short wave-lengths, then all the properties could be accounted for. Such has been shown to be the fact.

Light is diffracted and broken up into spectra by a grating made by ruling with a diamond point a large number of parallel lines, either on glass or on other suitable material. Such a grating has from ten to forty thousand lines to the inch.

Since it is impossible to make a grating with a sufficient number of parallel lines to the inch to reflect the X-rays, provided they could be reflected, it was necessary to find a suitable substance in nature. In 1912 Laue conceived the idea that the regularly arranged particles of a freshly broken crystal ought to furnish just such a space grating. This was soon proved to be the case, by Bragg and others. In short, it has been definitely established that, when X-rays are caused to fall upon such a crystal, the rays are broken up into spectra by the spaces between the molecules making up the crystal. This means that X-rays, like light, are a series of vibrations, differing only in being of short wave-length.

This series of experiments is not only giving the desired result in regard to the X-rays, but is also giving us pictures of the internal structure of crystals. When X-rays, after reflection or refraction by a crystal, are allowed to impinge upon a photographic plate or fluorescent screen, patterns are produced whose form changes as the crystal is rotated in different directions; and, from a study of the patterns thus obtained with the crystals in different positions, *the arrangement of the atoms and the molecules within the crystals can be determined.*

9. The Mass of Electrons.—As previously stated, the term electron was applied by G. J. Stoney to designate the unit or atomic charge of electricity, and it is now almost universally applied to the sub-atomic particles that stream from the negative electrode when a discharge is passing through an attenuated gas. No difference can be detected in the corpuscles from the different gases, and hence it is inferred that *electrons are common constituents of all gases.*

It has likewise been stated that when X-rays pass through a gas, the gas becomes ionized. Now, there is evidence to show that these gaseous ions, as well as the electrons liberated from them, all possess the random motion of the molecules of the gas, that is, they move hither and thither, colliding with anything which may come in their path. If tiny droplets of some liquid, such as oil, are introduced into such an ionized gas, they will sooner or later collide with one of these swiftly moving ions, or the free electrons, and will absorb and hold them. Each droplet, after it has collided with the ion or electron, will thereby acquire a charge of electricity exactly equal to that possessed by the absorbed ion. It is evident that, by observing the behavior of the charged droplet in an electric field, the magnitude of the elementary charge of electricity can be measured. This has been done with remarkable accuracy by Professor Millikan of the University of Chicago. He has given us indisputable proof that electricity is corpuscular and that the electron is the unit quantity.

The sketch below gives a fairly accurate idea of the apparatus used by Professor Millikan.

By means of an atomizer, *A*, a finely divided spray of oil, or of some other liquid, is introduced into chamber *D*, and one of the tiny droplets eventually falls through the aperture, *p*, into the space between the two condenser plates, *M* and *N*. The hole, *p*, is then closed, and the air between the plates is ionized by a beam of X-rays from *X*. The arc lamp, *a*, serves to illuminate the drop through the

windows *g* and *c*, and its motion is observed through a telescope, at a third window not shown in the figure. The eyepiece of the telescope is fitted with an engraved glass scale, which enables one to measure the motion of the droplet with great accuracy. When the droplet collides with one of the ions of the gas in the condenser chamber, it acquires an electric charge, and can then be caused to move upward or downward at will by charging the plates, *M* and *N*, with electric charges of the desired signs, by means of the switch, *S*. Any change in the speed of the droplet from that due to the force of gravitation is an indication of a change in the charge that it carries. Therefore, since the speed with which the droplet moves is directly proportional to the magnitude of its charge, it is evident that, by measuring this change in the speed of the droplet just after the capture of an ion and comparing it with the speed of the droplet when falling under the influence of gravitation, the magnitude of the charge of the captured ion can be calculated. Millikan found that the charge on the droplet was invariably either identical
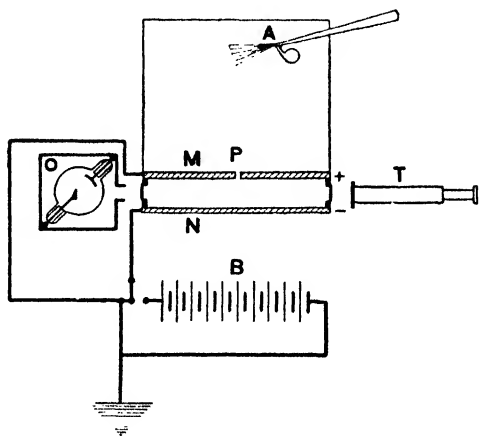
Fig. 2.—Diagram of Millikan's Apparatus for the Determination of the Mass of Electrons.

with the smallest charge ever found or an exact integral multiple of it, depending, of course, on the number of electrons attached. Thousands of observations were made, and not one exception to this rule was found. The actual value of the electronic charge was found to be 4.774×10^{-10} electrostatic units ($= 1.59 \times 10^{-19}$ coulombs), with a probable error of 0.2 per cent. Furthermore, the ratio of the charge to the mass (e/m) was found by Millikan's experiments to be 1845 times as great as in the case of the hydrogen ion. Now, since the charge is of the same magnitude, although of opposite sign, the mass of the electron must be $\frac{1}{1845}$ that of the hydrogen atom.

Sources of electrons other than the cathode rays are now available, and the nature of electrons has been satisfactorily proved to be independent of their source. For example, by means of experiments, which

need not be described here, it has been shown that in all probability the electric charge on an ion formed in the process of electrolysis is the same as the charge of an electron. Also, many solid bodies, especially metals, under the influence of high temperatures, give off electrons. It seems to be proved, therefore, beyond any reasonable doubt, that electricity exists in corpuscular form, and that the electron is the unit corpuscle, also that these corpuscles of electricity are component parts of all atoms. Moreover, it is now generally conceded that the electron theory explains, more clearly than any other theory that has yet been advanced, the splitting up of molecules in certain solvents into separate constituents (ions) carrying electrical charges. (See Chapter VIII.)

10. Positive or Canal Rays.—After the atom of negative electricity had been isolated, in the form of cathode rays, the next suggestion was that, possibly, corresponding positive electric particles might be discharged from the anode in a vacuum tube. Researches along this line were also successful. When a perforated cathode is employed in a vacuum tube for producing cathode rays, a stream of violet light passes through the perforations or canals and emerges behind the cathode on the side remote from the anode; hence, these rays are called canal rays. The canal rays have been investigated in the same way that the cathode rays were studied. The results indicate that the canal rays are streams of particles the majority of which are positively electrified; hence the term *positive rays* is replacing the older name, canal rays. The streaming particles travel in straight lines and produce phosphorescence (usually violet) when they impinge upon glass, etc. The speed of the positive rays is usually much less than that of the electrons; and they are not so sensitive to magnetic influences. Measurements similar to those used for the cathode rays show that the positively charged particles must be of atomic dimensions, and in no case is the mass of the positive corpuscle less than that of the hydrogen atom. Remembering that, so far as we can tell, all electrons are the same, and have a mass 1845 times less than that of the hydrogen atom, while the mass of the positively charged particles depends upon the nature of the gas and is virtually the same as that of the atom from which it is derived, it is probable that *when a gas is ionized, one or more negatively charged particles, electrons, are expelled from the atom, and the corresponding positively charged nucleus remains.*

11. Naturally Radioactive Substances.—The discovery of the X-rays and the study of their properties led to highly important discoveries, such as that of the existence of naturally radioactive substances. It had been shown by Roentgen that the portion of the vacuum tube where the X-rays originate becomes brilliantly fluorescent, and this led to the assumption that this fluorescence might be the cause of the new type of radiation. The study of those substances which fluoresce under the stimulus of the sun's rays was also undertaken. Among those who became interested in these phenomena was Henri Becquerel.[1] He undertook the examination of fluorescent substances to see if any of them gave off a radiation at all analogous to the X-ray. Among the substances that he studied was the double sulphate of potassium and uranium. This substance, after exposure to sunlight, was found to produce an impression on a photographic plate wrapped in black paper to cut off all ordinary light. The radiations given off by the salt of uranium could pass through thin sheets of metal and still affect the photographic plate. It later developed that all the salts of uranium would produce the effect, both those that were fluorescent, and those that were not. Becquerel also showed that stimulation by sunlight was not a necessary factor in the phenomenon, for the radiation was given off even when the uranium salts had not been exposed to sunlight. In fact, it has since developed that salts of uranium that have been kept in the dark for years are just as active as those that have recently been exposed to bright sunlight. The energy of the uranium radiation is therefore intrinsic in the uranium, and does not come from any external source. Other substances have been found that emit similar radiations spontaneously, and the phenomenon is known as *radioactivity*.

At the suggestion of Becquerel, Mme. Curie (Marie Sklodowske Curie), Professor of Physics at the Sorbonne, Paris, studied uranium minerals and found them to be more active than the uranium salts. Furthermore, it was found that the residues of uranium minerals from which all the uranium had been removed were much more radioactive than uranium itself. This led Mme. Curie to conclude that the activity was due to the presence in the natural minerals of a small quantity of some strongly radioactive substance which was neither uranium nor any known substance, but a new element.

[1] Compt. rend., **122**, 689 and 762, 1896.

With the assistance of her husband, Mme. Curie undertook the task of separating this unknown substance from pitchblende, a mineral which contains, in addition to uranium, a large number of other elements in small quantities. The first important discovery that was made was that of the existence of an element resembling bismuth, which Mme. Curie called polonium in honor of her native land, Poland. Later she isolated another radioactive element which resembled barium in many of its properties. This element she called *radium*. This new element is found in very minute quantities. One ton of pitchblende yields about 0.1 gram of radium chloride, which is the salt usually isolated. The salts of radium resemble the corresponding salts of barium in most of their chemical properties. Radium, therefore, appears to be a member of the family alkaline earths. Furthermore, the spectrum of radium resembles the spectra of the alkaline earths, which consists of strong lines and also bands. Next to uranium, radium has the highest atomic weight. The latest figure is 225.95.

Besides radium, many other elements have been found to be radioactive, but none are as active as radium, which is nearly two million times as active as an equal amount of pure uranium salt.

12. Nature of Radioactive Rays.—All the radioactive substances possess the power of acting on a photographic plate and of ionizing any gas in their immediate neighborhood. The intensity of the radiations may be compared by means of their photographic or electric action, and, in the case of strongly radioactive substances, by the power they possess of lighting up a phosphorescent screen. Such effects, however, are not sufficient to differentiate them from cathode rays or from X-rays, although each of these effects establishes the intensity of the radiations.

The first noteworthy evidence in regard to the true character of the radiations was presented by Sir Ernest Rutherford, while Professor of Physics at the University of Manchester. He examined the relative absorption of the rays by solids and gases, and observed their behavior when subjected to the action of magnetic or electric fields; he found that three distinct types of radiations were emitted from most radioactive substances. One of these rays was mildly deflected in a magnetic field and had moderate penetrating power; another was easily deflected in a magnetic field, and had fairly powerful penetrating power; the third ray was not deflected at all in a magnetic field, but had great penetrating power. Rutherford named these three distinct

types of radiations, the alpha, beta, and gamma rays, respectively. Roughly speaking, the beta rays are 100 times more penetrating than the alpha rays, while the gamma rays are 100 times more penetrating than the beta rays.

The general characteristics of the three kinds of rays may be briefly summarized as follows:

The *alpha rays* are given off by all radioactive substances. They are somewhat deflected in the magnetic field. They have very small penetrating power, being easily absorbed by very thin layers of matter. They have great power to ionize a gas, rendering it a conductor. The rays ionize to about one hundred times the extent of the *beta* and *gamma* rays together. They have but little effect on a photographic plate, but they produce phosphorescence in certain substances, especially zinc sulphide. The alpha particle has a mass of the order of magnitude about four times that of the hydrogen atom. These properties are almost all identical with those of the *canal* or *positive rays*. Recent researches of Rutherford have given proof that cannot be refuted *that the alpha rays are nothing but helium*, a gaseous element that was first found by spectrum analysis in the sun. The alpha particles carry two positive charges, and move with a velocity about one-tenth that of light.

The *beta rays* are given off by all radioactive substances, with the exception of polonium. They are easily deflected in the magnetic field; they affect the photographic plate; they excite phosphorescence; they ionize a gas; they are negatively charged particles; they have moderate power to penetrate matter; they have a mass about one-eighteen-hundredth that of hydrogen; they have a velocity about one-half that of light. Now, all these properties are identical with those of the *cathode rays*, except the velocity at which they travel. The *beta rays* are, therefore, considered to be swiftly moving *electrons*.

The *gamma rays* exist where the *beta rays* exist. They are not deflected in a magnetic field at all; they have great penetrating power; in fact, their penetrating power is enough, after passing through a foot of iron, to be detected by an electroscope; they have considerable power to affect a photographic plate, much greater than that of the alpha or beta rays; they have considerable power to ionize a gas; they excite phosphorescence; *they are not material, like the alpha or beta rays*. They are, in fact, so similar to the X-rays in their properties, that the most probable theory as to their nature is

that they are a very penetrating form of *X-rays*, produced by the *beta rays*.

It must not be thought, however, that every radioactive substance emits all three of these rays at once, for this is not the case. Only a few elements emit all three rays; others emit alpha rays alone; and still others emit beta and gamma rays. Most of the more common radioactive elements, such as uranium, thorium, and ionium, emit alpha rays only. The beta and gamma rays may be emitted from some of their disintegration products. Actinium itself is almost rayless, but it changes into radioactive emanations, some of which emit alpha, beta, and gamma rays; some, alpha rays only; and others, beta and gamma rays. All three types of rays, however, are continuously emitted by radium compounds in their normal state.

13. Other Properties of Radioactive Substances.—Besides their powers of acting on a photographic plate, and of ionizing gases, the radiations from radiating substances are able to produce marked chemical and physical effects in a variety of substances. These effects are mostly due to the alpha and beta rays. The similarity, if not identity, of the beta rays with the cathode rays implies that they will produce the same effects as those produced by the cathode rays. The following are some of the distinctive characteristics of radioactive substances:

(a) *Luminosity of Radium Compounds.*—All radium compounds are spontaneously luminous, and the self-luminosity seems to come from the entire mass of the substance, and not merely from the surface. Even solutions of radium salts are slightly self-luminous. In supersaturated solutions the crystals that separate out are more luminous than the solution and can be seen by the light which they emit.

Very active preparations of actinium show a similar luminosity, the intensity varying with the nature of the impurities present. Crystals of uranium nitrate were observed by Becquerel to be slightly luminous.

(b) *The Phosphorescence Produced by Radium Salts.*—We have already seen that the ability to excite phosphorescence in certain substances is common to all radiations, but radium shows some distinctive effects. For instance, the action of radium is not the same as that of the X-rays. Certain substances phosphoresce when exposed to the X-rays, but do not do so in the presence of radium, and *vice versa*. The action of radium is more nearly like that of ultra-violet light. Barium platinocyanide and zinc sulphide are very sensitive to radium

salts. The platinum salt shows phosphorescence when placed 2 meters away from the active radium. These substances have been found to be especially sensitive to the alpha rays, although the other two rays are mildly effective. This has been proved by the cutting off of the alpha particles by means of a screen of thin metal or paper.

(c) *Intra-atomic Energy.*—When coal is burned, heat energy is evolved through the attraction existing between the atoms of carbon (coal) and the oxygen of the air. This is in accord with the general rule that whenever there are strong attractions or repulsions between different atoms energy is usually evolved in measurable amounts. In the case of radioactive substances, however, energy is evolved from forces acting within the substance, and not from forces acting between atoms, as is the case with all chemical energy. Moreover, these intra-atomic forces are great as compared with the extra-atomic forces of chemical action. For example, pure radium chloride, without cessation and for an almost indefinite period, evolves enough heat to maintain itself at a constant temperature of 1.5° C. above other objects in the room. A gram gives off 118 calories every hour, an amount sufficient to raise a gram of ice water to the boiling-point. It has been estimated that a gram of radium will continue radioactive for about 2500 years, and it therefore follows that a gram of radium gives off energy equivalent to that obtained by burning nine-tenths of a ton of coal during the period of radioactivity. Otherwise expressed, the *intra-atomic* energy of radium is 250,000 times as great as the extra-atomic energy given by burning an equal amount of coal. Furthermore, the rate of emission is unlike that of chemical forces, in that it is unaltered by change of temperature. Even cooling by immersion in liquid air has no effect.

Several theories have been advanced to account for the production of the heat that is continuously being liberated from radium. The most satisfactory, perhaps, is the one advanced by Lodge.[2] He contends that it is only the surface alpha particles that are shot off from the radioactive substance, whereas the alpha particles that are shot off underneath the surface of the salt do not escape, but strike the solid salt above them and are stopped. The energy of motion of the alpha particle would thus be converted into heat energy. Since the mass of the alpha particles is of atomic dimensions (being helium) and the velocity is about one-tenth that of light, it is evident that the kinetic energy (and therefore heat) produced would be great. This theory

[2] Nat. **67**, 511, 1903.

explains the source of the energy as we perceive it, but it does not explain where the alpha particles acquire the great velocity with which they are shot off from the radium. According to H. Jones [3] this kinetic energy comes from the intrinsic energy of the radium atom. To quote his own words: "We can scarcely conceive of particles at rest in a molecule being shot off with such velocities. The particles in the molecule or atom of radium—the electrons—must be moving with very high velocities, and when a particle in its motion gets beyond the control of the attractions of the remaining particles of the system, it flies off. This is true of the positively charged alpha particles, and also of the negatively charged beta particles. The kinetic energy of these particles is then something inherent in the atom of radium. This we call *intrinsic energy*. It is obvious that this is the real source of the heat liberated by radium. The astonishing feature is the amount of the intrinsic energy contained in the atoms of radium."

(*d*) *Physiological Action of Radium.*—Under the influence of radium, nutrition is profoundly modified, the development of growing animals is arrested, and, after prolonged exposure, they often die. When salts of radium are brought near the skin, it is first inflamed and reddened, and then blistered. These burns are slow to heal and are very painful. If a capsule containing some radium is held in the hand a few minutes, the skin will become inflamed, and, after remaining so for a couple of weeks, will peel off. The hand may be painful for months. This effect is brought about mainly by the alpha and beta rays. The gamma rays have been found to be beneficial to cancerous tissue, but, since the alpha and beta rays are even more irritating to pathological tissues than to normal tissues, they must be kept away from the tissue. This is done by keeping the radium salt in a lead tube which holds back the alpha and beta rays, but not the gamma rays. Since the gamma rays are more penetrating than ordinary X-rays, their curative properties are greater.

Another physiological action of radium has been thoroughly studied by Himstedt and Nagel.[4] If a radium compound is brought near to the closed eye in a darkened room, a sensation of light is produced. This is due to the phosphorescence produced within the eye itself by the radium, the lens and the retina being strongly phosphorescent under the action of the beta and gamma rays. The sensation is

[3] Elec. Nat. of Matter, p. 116, 1915.
[4] Ann. d. Phys., 4, 537, 1901.

experienced even by the blind, if the retina has not been destroyed.

(c) *Emanations from Radioactive Substances.*—In addition to the alpha, beta, and gamma rays, and the heat, certain radioactive elements, such as thorium, radium, and actinium, give off what Rutherford calls an *emanation.* This substance behaves as if it were a radioactive gas. It diffuses through certain bodies, which are known to remove ions and rays, and it can be condensed to a liquid at low temperatures. The first compound that was observed to yield an emanation was thorium. Mme. Curie and others, in studying the radiations from thorium, observed that the radioactivity of this compound was not constant if air currents were present, but that if the active thorium were placed in a closed vessel constant results were obtainable. On the other hand, if a current of air was drawn through the closed vessel the major portion of the activity seemed to be present in the air and not in the vessel containing the thorium. Rutherford[5] took up the study of this peculiar phenomenon, and soon established that the active thorium gave off a radioactive gas, which he called " thorium emanation." The quantity of this gas that was given off was always very small, but its activity was very pronounced, and was found to consist exclusively of alpha rays.

Radium and actinium were also found to emit radioactive gases, or " emanations." These emanations likewise emit only alpha rays.

The radium emanation is the one that has been most studied, since it is obtainable in greatest amount. In 1911, Ramsay and Gray,[6] by using a special quartz micro-balance sensitive to one-millionth of a milligram, were able to determine the mass of a known volume of pure radium emanation, and found its molecular weight to be 222. It was found to be inert, as far as ordinary chemical activity was concerned. This fact, taken together with the atomic weight of the emanation, served to determine its position in the group of non-valent elements in the periodic table. Ramsay proposed the name of " niton " for the new element. It is the heaviest known gas, having a density 111 times as great as that of hydrogen. It has been condensed by liquid air to the liquid state. The liquid has a specific gravity of 5.7 and boils at about $-62°$ C. At a still lower temperature the liquid solidifies. The solid melts at about $-71°$ C. The emanation, if left by itself, slowly

[5] Phil. Mag., **49**, 1900.
[6] Proc. Roy. Soc., **84**, 536, 1911.

disappears. At the end of four days, only about half the original quantity remains. How much more unstable this substance is than its parent, radium, is evidenced by the fact that radium takes about 1300 years to transmute itself by half. The radium emanation (niton) decomposes continuously and spontaneously into a radioactive solid and alpha particles. The alpha particles were later proved to be pure helium gas. This latter observation is most important, since it has a bearing, as we shall see later, on the whole subject of the composition of matter. The following experiment of Rutherford and Royds,[7] gives conclusive evidence that each alpha particle, regardless of its source, is an atom of helium carrying two positive charges. A large quantity of radium emanation was compressed into a thin-walled glass tube surrounded by a vacuum jacket. Both vessels were gas-tight. The walls of the inner tube were thin enough to allow the alpha particles to pass through, but sufficiently strong to withstand atmospheric pressure. The outer jacket would not permit the passage of the alpha particles. It carried a spectrum to enable the operator to examine the gas spectroscopically. The inner tube was filled with radium emanation, and the outer tube was exhausted. After two days, the spectrum showed the yellow lines of helium, and after six days the whole helium spectrum was observed. To show that the helium was not derived by diffusion from the inner tube, the emanation was withdrawn and helium substituted. No trace of helium could be detected in the outer vessel after standing for many days. This and many similar experiments proved that radium passes into helium *via* the emanation. Furthermore, the alpha rays are streams of positively charged particles, and each alpha particle, which carries two unit charges, becomes a helium atom when its charge is neutralized.

14. The Rate of Production of Helium from Radium.—Rutherford and Boltwood carefully collected and measured all the helium from a given sample of radium during a period of four months. The amount obtained was 39 cu. mm. of helium gas from a gram of radium per year.

15. Counting the Alpha Particles.—If an alpha particle strikes a screen covered with zinc sulphide (a spinthoriscope), it produces a small flash of light which can be observed with a microscope. To estimate the number of alpha particles that a given quantity of radioactive matter gives off in a given period of time, it is only necessary to count the number of flashes produced on such a screen. An electrical method

[7] Phil. Mag., VI, 17, 281, 1909.

for counting the alpha particles has also been devised by Rutherford and Geiger.[8] In their experiment, the source of the alpha particles was a small disc which had been exposed to the emanation for some time. This disc was placed in a long evacuated glass cylinder, called the *firing chamber*, at a definite distance from an opening of known cross-section. This opening consisted of a thin mica plate through which the alpha particles could easily pass into the detecting chamber, which was fitted with two electrodes charged at a high potential and connected with a delicate electrometer, an instrument for detecting the passage of a current. Whenever an alpha particle entered the detecting chamber, it ionized the air, and a momentary current of electricity flowed from one of the electrodes to the other, thereby producing a sudden deflection of the needle of the electrometer. The experimenter had only to count the number of throws of the needle occurring in a given time to know the total number of alpha particles that passed through the ionization chamber in that time. Knowing the distance of the source of the radiations from the opening, together with the area of the opening, one could estimate the total number of alpha particles emitted by the radioactive substance in a given time, since they are emitted uniformly in all directions. As a result of a number of experiments, Rutherford and Geiger found that the number of alpha particles emitted per second for one gram of radium is 3.4×10^{10}. These results were consistent with those obtained by counting the number of scintillations produced when a zinc sulphide screen was used as a target.

16. Calculation of Avogadro's Number from Radioactive Data.—By combining Rutherford and Geiger's value for the rate of emission of alpha particles by radium with Rutherford and Boltwood's value for the rate of production of helium from radium, Avogadro's number can be calculated. The result, which is 61.5×10^{22}, agrees very well with the values obtained by methods based on other principles.

17. The Law of Decay of Radioactive Substances.—As stated before, the rate of decay of a radioactive element, from all present indications is independent of all external conditions. The law of decay is that, for a given element, the number of atoms which decompose in a given period of time is proportional to the number in existence at that time. In other words, if we should start with a gram of radium A (this is the name given to the radioactive salt left after radium emanation

emits helium), it would be half gone at the end of 4.3 minutes. In 4.3 minutes more, half of the remainder, or one-quarter of the original amount, in addition, would have broken up, and so on. It is evident that a law of change of this type would theoretically never lead to the total disintegration of any given quantity of the substance. For this reason, the " life " of a radioactive element is stated in terms of " half periods," i.e., the time required for one-half of a given quantity of the active element to decompose.

These half-periods vary from about twenty-five billion years in the case of thorium to only one-billionth of a second in the case of radium C. Of course, the rate of decomposition of radioactive elements with very long or very short " lives " has not been determined directly, but is merely calculated from observations of the distance that the alpha particles are able to penetrate air. For it happens that the shorter the " life " period of an element the more energetically are the particles hurled off from the atoms when they " explode."

18. The Disintegration Theory.—The only theory thus far proposed that accounts at all satisfactorily for the phenomena discovered about radioactive elements, such as the emission of alpha, beta, and gamma rays, and the generation and decay of emanations and active deposits, etc., is that advanced by Rutherford and Soddy.[9] According to this theory, the atoms of the radio-elements are assumed to undergo spontaneous disintegration, each atom passing through a series of well-defined changes. A definite number of the atoms of any radioactive element become unstable in a given time, and may then throw off an alpha particle (helium) or a beta particle (electron), leaving behind another element with more or less different properties. The expulsion of an alpha particle, i.e., a helium atom with an atomic weight of 4, results in the production of a new element possessing quite different physical and chemical properties. For example, such an element has an atomic weight 4 less than the parent element. On the other hand, if the particle thrown off is a beta particle (electron) the atomic weight will remain unchanged, since an electron has such an infinitesimal atomic mass, only $\frac{1}{1845}$ the mass of the hydrogen atom. Once the process of disintegration sets in, it continues successively from one radio-element to the next, and continues through a long series of definite stages, until finally a stable compound is produced and further disintegration ceases. In considering the modes of transformation of

[9] Phil. Mag., 5, 576, 1903.

uranium, thorium, and actinium, and the chemical properties of their successive products, Rutherford [10] says:

" Each element gives rise to a radioactive emanation which behaves like a monatomic gas of the helium-argon group, and each emanation in turn gives rise to a similar series of products included under the name ' active deposit.' In these groups of products, the similarity is very close, not only in the nature of the transformations but also in the chemical characters of corresponding products. Each emanation is transformed with the emission of an alpha particle and gives rise to a non-volatile product of short life known as radium A, thorium A, and actinium A, respectively. Each of the A products breaks up with the expulsion of an alpha particle, giving rise to a B product. Each of the B products has the longest life of the group, and is transformed into a C product. There is evidence that in all three cases the C products break up in an abnormal way. Radium C emits alpha, beta, and gamma rays, and gives rise to radium D and a branch product called radium C_2 of short life, which emits beta rays. Thorium C breaks up in a complicated way, emitting two groups of alpha rays and also beta rays. There is also some evidence, though of a very definite character, which indicates that the transformation of actinium C is abnormal. The products thorium D and actinium D are very similar, both emitting beta and gamma rays, and having a short life compared with the B products. Thorium D and actinium D differ in some respects from radium D. It seems probable that the product of radium analogous to thorium D is included under radium C, for it is quite possible that the beta and gamma rays from the latter do not accompany the expulsion of alpha rays. . . . The general analogy between the modes of disintegration of the first four products of the emanation is thus very striking. The similarity extends to the chemical nature of the products, for it has been shown that the B and C products of the series show almost identical chemical behavior. On the other hand, the series of products radium $D + E + F$ has no counterpart in the thorium or actinium series."

Taking the three series together, there are nearly forty radioactive elements, each of which is endowed with specific properties and with periods of existence varying from billions of years on the one hand, to a small fraction of a second on the other. The table shown in Fig. 3 gives the three disintegration series as arranged by Soddy.

[10] Radioactive Substances and their Radiations.
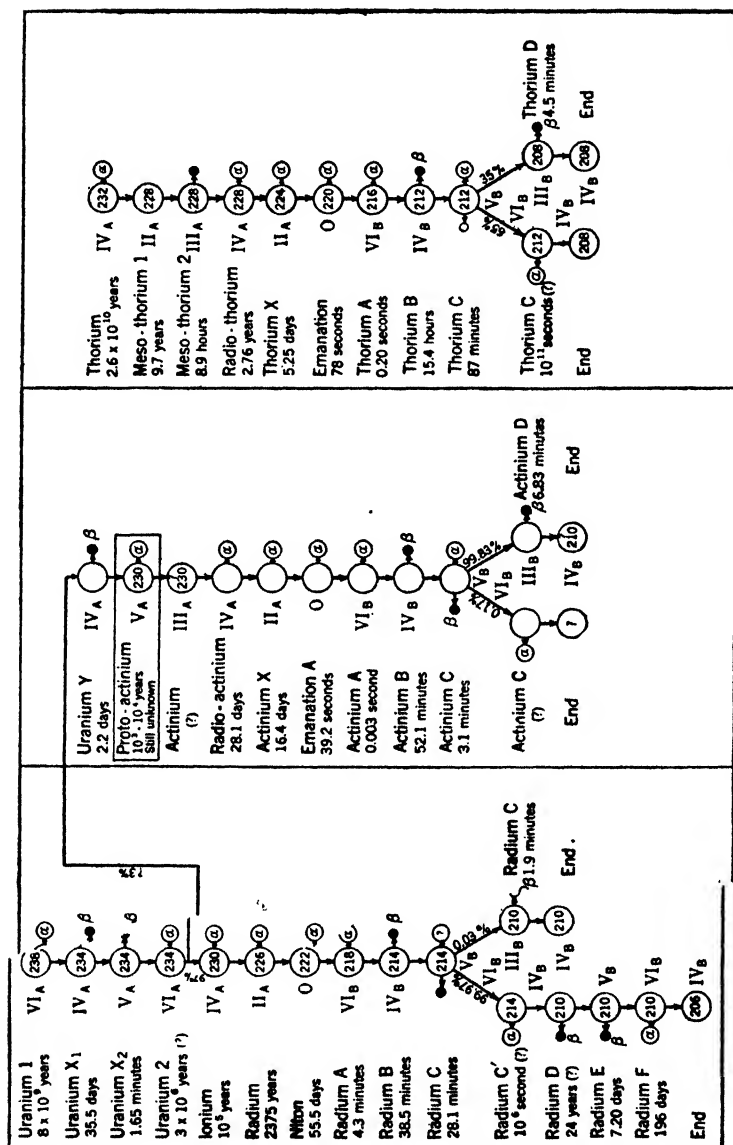
FIG. 3.—Table Showing Disintegration Series.

The numbers within the circles denote the atomic weights of the elements, while the small circles and dots at the right of the larger circles indicate the character of the radiation given out at each stage of the disintegration process.   The average life, $\frac{1}{\lambda}$, of each element in the series, is given below the name of the element.

19. The Chemical Law of Radioactive Change and the Discovery of " Isotopes " and " Isobars."—Up to about 1913, the relation of the radio-elements to the periodic system remained an unsolved problem, but about this time Fajans,[11] Soddy,[12] and others brought forward an important generalization with regard to the changes in the chemical properties which result from successive steps in a radioactive disintegration.   This generalization has an important bearing upon the modern theory of atomic structure.   It may be briefly expressed in the following terms: *A radioactive element on losing an alpha particle goes back two places in the periodic table; on losing a beta particle it goes forward one place.*   This generalization not only agrees with our present theory of atomic structure, but is a necessary consequence of this theory.   Soddy says:

" It has been possible to predict from the generalization many new details, and these have been almost immediately and conclusively established by experiment. . . . In consequence we can now write the three disintegration series across the Periodic Table, as in Fig. 4, so that the series fits into the proper place in the Periodic Table; and, so far as can be ascertained, all the elements occupying the same place are chemically non-separable and identical.   The diagram is interesting, not only as establishing a fair presumption that the whole of these new changes have now been disentangled, but even more interesting because of the new light it sheds on what the Periodic Law means, and what the chemical elements themselves really are.   The chemical non-separability and identity of different radio-elements, first definitely drawn attention to in the case of radium and mesothorium I[6], and now shown to be the general characteristic of the radio-elements, becomes something more than the expression of experimental failures to separate and distinguish."

The most notable feature of this table is the fact that a single place in the periodic system is filled by a number of different elements.   For

---

[11] Physik. Z., **14**, 131, 1913.
[12] Chem. News, **107**, 97, 1913.

example, the place occupied by lead in the table is also filled by Ra-B, Ra-D, Th-B, and Ac-B, together with five other unnamed elements which are the stable end products of the different disintegration series. These end products seem to be related to lead in every case, that is, they are chemically indistinguishable from lead, and cannot be separated from it by any known chemical means, and they are spectroscopically indistinguishable. We note, however, that four different
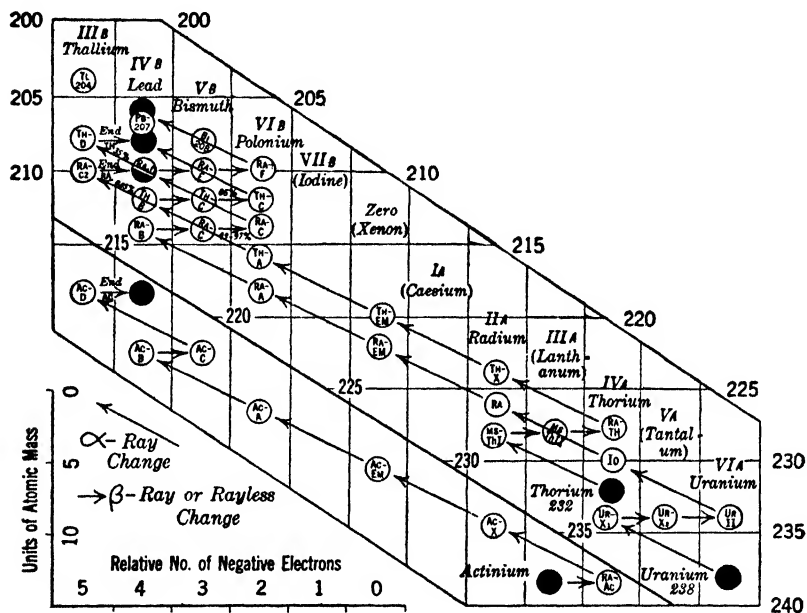
FIG. 4.—Radio-elements and the Periodic Law. All elements in the same place in the periodic table are chemically non-separable and (probably) indistinguishable. (After Soddy.)

atomic weights are involved: 207.2 for lead, and, for the different radioactive end products, 206, 208, and 210.

Soddy proposed the general name of " isotopes " for elementary forms of matter which appear to be identical in chemical properties, but which differ in atomic weight and in properties depending on atomic weight. Recent investigations show that we have many examples of isotopes outside of the radioactive series, e.g., Cl, Zn, Hg, Ni, etc. (see Art. 21).

Just as we can have elements of the same chemical properties but

with different atomic weights, so it is likewise possible to have elements possessing different chemical properties with the same atomic weight. These elements have been called "isobars" by Stewart.[13]   Any product of radioactive change due to the loss of a beta particle (which has a negligible mass) is an isobar of the parent element, for, without change of mass, it has shifted its position in the periodic table and so changed its chemical properties. So far, no isobar has been discovered outside of the radioactive series. Still, we have reason to believe that they do exist.

20. Are all Elements Radioactive?—There is reason to believe, but no absolute proof, that all elements are undergoing radioactive disintegration. This belief has been brought about largely by the discovery of a number of feebly radioactive substances, such as potassium and rubidium. Potassium has about one-thousandth the radioactivity of uranium.   The radioactivity of the so-called non-radioactive substances is supposed to be taking place too slowly to be detectable.

21. The Determination of the Mass of an Atom: Confirmation of the Isotope Concept.—In the creation of the concept of "isotopes," Soddy advanced the idea that the occurrence of isotopes is not confined to the radioactive elements, but that many of the common elements might also be isotopic mixtures. It is evident that, if such is the case, a "mixed element," which has an uneven number for its atomic weight, might be made up of isotopic elements having whole numbers for their atomic weights. Experimental proof as to whether an ordinary non-radioactive element consisted of a definite mixture of isotopes was not easily attainable, owing to the similarity of every one of the properties which might be used for separating the different substances. The atomic weight, being the only measurable factor, had to form the basis of the methods which were devised for separating the isotopes. The older methods, however, which were used in estimating the atomic weight of an element, were not practicable in this case. It was not until the so-called "positive-ray method" was devised that satisfactory results were obtained.   This method, devised by Sir J. J. Thomson, depends upon the fact that, in the cathode-ray tube, electrons, which are shot out from the cathode under the influence of high potential, ionize the gas in their path, that is, by collisions with the atoms of the gas they dislodge one or more electrons from these atoms, thus leaving them positively charged.   These positively charged atoms are imme-

[13] Phil. Mag., **36**, 326, 1918.

diately subjected to a powerful attraction by the negatively charged cathode, and move towards it with great speed. If a tube with a perforated cathode be used, these positively charged atoms will pass through the perforations into the space behind the cathode, where their properties may be studied. If a quantity of these rays, called *positive rays* by Thomson, are subjected simultaneously to powerful electric and magnetic fields, the particles will be deflected through an angle which will depend upon the ratio of the charge on the particle to its mass, e/m, and if they are then allowed to impinge upon a photographic plate, an image will be produced at a position which corresponds to the magnitude and direction of such deflection. While the method just described has proved of great value for a general survey of masses and velocities, still it lacks precision.

The defects of Thomson's method have been largely overcome through the efforts of F. W. Aston,[14] who has recently perfected an apparatus for positive-ray analysis, which he calls a "mass spectroscope."

The spectrograph consists of a glass discharge tube of about 20-cm. diameter provided with an aluminium anode in a side tube at one end of the discharge chamber. The discharge is maintained by a large induction coil furnishing 0.5–1 milli-amp. at 20,000–50,000 volts. The glass tube is protected from the cathode-ray bombardment by an anticathode of silica which has the advantages of infusibility and of emitting no disturbing X-radiation. The discharge bulb is large in order to faciliate conduction at very low gas pressure. The cathode consists of two parallel slits placed in the neck of the chamber opposite the anode. They are of aluminium; the apertures are 0.05 mm. wide and 2 mm. long. Their object is to furnish a narrow uniform beam of positive rays, and their correct alignment is important. The space between the two cathode slits is highly evacuated by a side tube of charcoal immersed in liquid air. The positive rays after passing the slits enter the electrical field of 200–500 volts between brass electrodes 2.8 mm. apart. After passing the electrical field the ions reach the magnetic field between pole pieces of a large Du Bois magnet of 2500 turns, the faces of which are 8 cm. in diameter and 3 mm. apart. The rays then emerge into the photographic chamber which is provided with an arrangement for shifting the plate without opening. The position of the deflected

---

[14] Phil. Mag., **38**, 709, 1919; *ibid.*, **39**, 611 1920. (See also Aston's monograph, "Isotopes," 1922.)

beam recorded on the plate is read by means of a comparator giving both dimensions, from which the e/m value of the particles constituting the beam can be calculated.

Measurements made with this apparatus rendered it possible to determine the atomic weight of individual atoms with an accuracy of about 0.1 per cent.

The investigations performed by Aston on a large number of elements gave astounding confirmation of the hypothesis of "mixed elements." The following table given by Aston summarizes the results obtained up to 1924 by the postitive-ray method:

TABLE I

ISOTOPES OF NON-RADIOACTIVE ELEMENTS *

Element	Mean Atomic Weight	Atomic Weights of the Isotopes in the Order of their Frequency of Occurrence	Element	Mean Atomic Weight	Atomic Weights of the Isotopes in the Order of their Frequency of Occurrence
H	1.008	1.008	As	74.96	75
He	4.00	4.00	Se	79.2	74, 76, 77, 78, 80, 82
Li	6.94	6, 7	Br	79.92	79, 81
Be	9.02	9	Kr	82.92	78, 80, 82, 83, 84, 86
B	10.82	10, 11	Rb	85.5	85, 87
C	12.00	12	Sr	87.63	86, 88
N	14.01	14	Zr	90.6	90, 92, 94, 96
O	16.00	16	Sn	118.70	116, 117, 118, 119, 120, 121, 122, 124
F	19.00	19			
Ne	20.20	20, 22	Sb	121.77	121, 123
Na	23.00	23	I	126.92	127
Mg	24.32	24, 25, 26	X	130.20	128, 129, 130, 131, 132, 134, 136, (124), (126)
Al	27.10	27			
Si	28.30	28, 29, 30	Cs	132.81	133
P	31.40	31	Ba	137.37	(136), (137), 138
S	32.07	32	La	139.00	139
Cl	35.46	35, 37	Ce	140.25	140, 142
Ar	39.90	36, 40	Pr	140.90	141
K	39.1	39, 41	Nd	144.30	142, 144, (145), 146
Ca	40.70	40, 44	Er	167.70	164–176 (impure sam.)
Fe	55.84	54, 56	Hg	200.60	197, 198, 199, 200, 202, 402
Ni	58.68	58, 60			
Zn	65.37	64, 66, 68, 70			

* Aston, Phil. Mag., 1923; and Nature, **113**, 856 (1924); *ibid.* **114**, 273 (1924).

The most noteworthy result brought out in this table is not that a number of elements really do consist of a mixture of isotopes, but that

the atomic weights of all the individual isotopes of which the elements are composed are *exactly whole numbers*, with one exception, namely *hydrogen*, which maintains the figure formerly obtained for it—1.008.

22. The Electrical Theory of Matter.—The facts covered in the foregoing paragraphs on the discharge of electricity through rarefied gases, the study of radioactive substances, the experiments on isotopes, and many others which could not be included have led to certain conclusions regarding the structure of atoms, the most important being centered around the assumption that the atoms of the elements are aggregations of atoms of negative and positive electricity.

The negative particle of electricity, or electron, has a mass that is only $\frac{1}{1845}$ that of the positive particle of electricity, or 0.00054 on the common scale of atomic weights (0 = 16).

The atom of positive electricity, now called *proton* (This name, which is taken from the Greek word πρῶτον, "first" = the primary substance, was suggested for it by Sir Ernest Rutherford at the Cardiff meeting of the British Association in 1920.), has been identified as being the hydrogen positive ray, i.e., the positively charged particle that remains when the neutral hydrogen atom loses an electron. The proof of the identity of the positive atom of electricity and the hydrogen "ion" is largely based on the fact that in none of the analyses of positive rays so far performed has a particle been discovered with a mass less than that of hydrogen. The charge on the proton is, of course, just equal to that of the negatively charged electron. This implies that the neutral atom of hydrogen consists of one electron and one proton. The mass of a proton is nearly identical with that of the neutral hydrogen atom, being 1.66×10^{-24} grams, or 1.0072 on the oxygen scale.

The arrangement of the atoms of electricity within the atoms of the elements has been considered by many prominent chemists and physicists, notably Thomson, Rutherford, Moseley, Nicholson, Bohr, Lewis, and Langmuir. We shall consider some of the more important points that have been brought out by these investigators.

Sir J. J. Thomson[15] advanced the theory that atoms consist of a system of electrons dispersed through a sphere of positive electricity, which is as large as the atom. The electrons are assumed to be at relatively great distances apart compared with the spaces actually occupied by them, like the planets in the solar system, and they are assumed to move with very high velocities. The electrons are so dis-

[15] Phil. Mag., **7**, 237, 1904.

tributed through the positive sphere as to be in dynamical equilibrium under the forces acting upon them. These are the attraction of positive electricity for negative electrons, and the repulsion of one electron by another.

Two of the objections that have been raised to the Thomson theory of atomic structure are especially worth mentioning. One is "That the atoms so formed would not be stable, for according to the electromagnetic theory, electrons in orbital motion must radiate energy, and hence at some time the atom would break up." The other objection is "That the positive spheres of Thomson would have electromagnetic inertia which would be negligible compared with that of even a single negative electron, leaving practically all the mass of the atom unaccounted for on this theory."

The first theory of the constitution of the atom with any experimental work as a basis is due to Sir Ernest Rutherford.[16] The atom model which he formulated has resulted in bringing about remarkable advances in both physics and chemistry, and it is now almost universally accepted as correct in fundamental principle. His theory is that an atom consists of a central, positively charged nucleus in which most of the mass resides, surrounded by a system of "planetary" electrons. The central nucleus contains all of the positive electricity of the atom, hence the weight of the atom and the radioactive properties are associated with the nucleus; but the chemical properties and spectrum, on the other hand, are properties of its planetary electrons.

In a neutral atom the positive charge on the nucleus must be equal to the sum of the negative charges on the planetary electrons.

In the later developments of this theory, the units of positive electricity are taken to be protons, hydrogen atoms which have lost one electron, and the mass of the atom is assumed to be largely due to the number of protons which it contains.

It may be recalled that "Prout's Hypothesis," pronounced in 1815, was that the atoms of the elements were made up of aggregations of atoms of hydrogen, and from the foregoing statement one might expect that his theory still holds. Such is not the case, however, because according to the electromagnetic theory the total mass of a body made up of negative and positive units will depend to a large extent upon the manner in which they are "packed" together, and upon the energy charge accompanying the formation of the atom. After making proper

[16] Phil. Mag., 41, 669, 1911.

allowance for this "packing effect," however, a very exact multiple relationship can be shown to hold for most of the elements.

23. Atomic Numbers and X-ray Spectra.—We have already learned that the atom as a whole is electrically neutral, since the excess positive charge carried by the nucleus must be equal to the number of electrons outside the nucleus. The number of electrons that are outside the nucleus is roughly equal to half the atomic weight. Now, if the elements are arranged in order of their atomic weights, starting with hydrogen, each element will have a position the number of which will be about half its atomic weight. This is especially true of the elements having a low atomic weight. This number is called the *atomic number* of the element, and it is now recognized to be more important and characteristic than the atomic weight. Beginning with hydrogen, the atomic numbers of the elements run from H = 1, He = 2, C = 6, O = 8, etc., up to U = 92. Van den Broek was the first to point out the possibility that the atomic number might be equal to the charge on the nucleus. This was subsequently proved by Moseley, Bohr, Chadwick, and others.

Moseley's[17] work on the X-ray spectra of the elements was very valuable in showing the relationship between the atomic number of an element and its X-ray spectrum. W. L. Bragg and his father, W. H. Bragg,[18] had shown that the X-rays undergo reflection at crystal surfaces, and later they invented an X-ray spectrometer. Following the lead of Bragg and Bragg, Moseley showed that X-ray spectra can also be used to determine the magnitude of the free nuclear charge of an element. He found that when X-rays were allowed to strike the surface of an element they were reflected off in such a way as to form a spectrum which could be photographed. Definite lines were always present, and the two most prominent of these he named the alpha and beta lines. The position of these lines was found to depend on the element used, and the lines shifted more and more towards the violet as the elements increased in atomic weight. In other words, Moseley found that each increase of one unit in atomic number—each advance that is, of one space in the periodic table—produced a definite constant increase in the *square root of the vibration rate* of the X-rays emitted by an element. It is thus possible to tell where an element belongs in the periodic table. After making a great many observations

[17] Phil. Mag., **26**, 1031, 1913.
[18] Proc. Roy. Soc., **88A**, 428, 1913.

and always finding the same relationship between the vibration rate and the atomic number, Moseley became convinced that the atomic number must represent the number of free charges on the nucleus.

From Moseley's law of atomic numbers we get a satisfactory explanation of the general law of radioactive transformations. This explanation was given on page 35 and will be repeated here: A radioactive element on losing an alpha particle goes back 2 places in the periodic table, and on losing a beta particle it goes forward 1 place. From Moseley's work, we modify this law as follows: If a radioactive element emits 1 alpha particle from its nucleus, it goes back 2 units in atomic number, which is the same as moving back 2 places in the periodic table. On the other hand, if it emits 1 beta particle, it must move forward 1 place, since the nucleus has thereby acquired 1 additional charge.

24. The Hydrogen-helium System of Atomic Structure.—A generalization similar to that just mentioned for the radioactive elements has been found by Harkins and Wilson[19] to hold true for the lighter elements which fail to show appreciable alpha-ray disintegration. They argued as follows: If a heavy radioactive element, through the loss of an alpha particle, which has an atomic mass of approximately 4 units of proton, is shifted 2 groups to the left in the periodic table and its atomic number decreased by 2, and if a similar relation holds true for the lighter elements, then, beginning with helium, the addition of 4 to the atomic weight ought to give the atomic weights of the elements belonging to the even-numbered groups of the periodic table, provided the change in mass due to variations in the manner and degree of "packing" in the nuclei of the successive elements are neglected. This rule they found to hold very closely for all of the elements having atomic weights less than 60. A similar rule was found to hold true for the elements of the odd-numbered groups in the periodic table, provided that the atom of lithium, the first member of the odd-numbered groups, was assumed to be made up of 3 hydrogen nuclei (proton) and 1 helium nucleus. The following table gives the results as calculated by Harkins and Wilson for the first three series of the periodic table:

The figures in this table that are marked "Theoretical" are computed on the basis of $H = 1$, while those marked "Determined" are the directly determined values on a basis of $O = 16$, or $H = 1.0078$.

[19] Proc. Nat. Acad., V, I, p. 276, 1915.

TABLE II

The Helium—H_3 System of Atomic Structure.   H = 1.0078

	0	1	2	3	4	5	6	7	8
Atomic Number..	2	3	4	5	6	7	8	9	
	He	Li	Be	B	C	N	O	F	
Series 2..........	He	He+H_3	2 He+H	2 He+H_3	3 He	3 He+H_2	4 He	4 He+H_3	
Theoretical......	4.00	7.00	9.0	11.0	12.00	14.00	16.00	19.00	
Determined......	4.00	6.94	9.1	11.0	12.00	14.01	16.00	19.00	
Atomic Number..	10	11	12	13	14	15	16	17	
	Ne	Na	Mg	Al	Si	P	S	Cl	
Series 3..........	5 He	5 He+H_3	6 He	6 He+H_3	7 He	7 He+H_3	8 He	8 He+H_3	
Theoretical......	20.0	23.0	24.00	27.0	28.0	31.00	32.00	35.00	
Determined......	20.0	23.0	24.32	27.1	28.3	31.02	32.07	35.46	
Atomic Number..	18	19	20	21	22	23	24	25	26 / 27
	A	K	Ca	Sc	Ti	V	Cr	Mn	Fe / Co
Series 4..........	10 He	9 He+H_3	10 He	11 He	12 He	12 He+H_3	13 He	13 He+H_3	14 He / 14 He+H_2
Theoretical......	40.0	39.00	40.00	44.0	48.0	51.0	52.0	55.00	56.00 / 59.00
Determined......	39.9	39.10	40.07	44.1	48.1	51.0	52.0	54.93	55.84 / 58.97

Increment from Series 2 to Series 3 = 4 He.   Increment from Series 3 to Series 4 = 5 He (4 He for K and Ca).   Increment from Series 4 to Series 5 = 6 He.

The remarkably close agreement between the two sets of figures is interpreted by Harkins and Wilson as being directly due to the *packing effect* in the formation of these elements from protons and electrons. They compute that the decrease is fairly constant, being about 0.77 per cent in the mass of the resulting atom. This decrease is believed to be almost entirely due to the formation of the helium atom.

This *hydrogen-helium hypothesis* of atomic structure offers a satisfactory explanation of many hitherto obscure facts concerning the nature of the elements, such as why the whole-number relationship among the atomic weights is much more exact on the basis of O = 16 than on the basis of H = 1. Oxygen is regarded as being composed of 4 helium atoms, which have an atomic weight of 4.00. (One helium atom equals 4 protons and 4 electrons, which would give an atomic weight of 4.032, but after deducting the 0.77 per cent for the *packing effect* we get an atomic weight of 4.00.)

Aston[20] explains the phenomenon of "packing" in the following terms:

"According to generally accepted views the proton and the electron possess *mass*, or what on the basis of the relativity theory is regarded as the same thing, *weight*, by virtue of the energy in the electromagnetical field which surrounds them. It can easily be shown on classical lines that if we give it a spherical form a charge e, spread uniformly over the surface of the sphere, will have a mass m when the radius a is such that $m = \dfrac{2}{3}\dfrac{e^2}{a}$; hence to give the electron its proper mass its charge must be compressed to a sphere of diameter about 3.8×10^{-13} cm. By the same argument, the proton will be nearly two thousand times smaller and have a diameter 2.06×10^{-16} cm. The extreme range of the diameter of atoms themselves is $1 - 5 \times 10^{-8}$ cm., so that it will at once be realized that the structure of an atom is an exceedingly open one, even more so than that of our solar system.

"Now it can be shown that if we bring two charges of opposite sign sufficiently close together their fields will affect each other in such a way that the mass of the system will be reduced. This effect is quite inappreciable for distances comparable with the diameter of the atom, but begins to make itself felt when the distance apart is of the order of the size of the electron itself as given above. The nucleus of the atom of an ordinary element (not hydrogen) contains both protons and elec-

[20] Isotopes, p. 100, 1922.

trons and is very small compared with the atom itself. Its dimensions can be roughly determined by actual experiment in the case of the heavy elements and are found to be so small that, even to get in the electrons alone, these would have to be packed very closely together. Such a nucleus will contain more protons than electrons, roughly twice as many, so that it may be regarded as practically certain that: *In the nuclei of normal atoms the packing of the electrons and protons is so close that the additive law of mass will not hold and the mass of the nucleus will be less than the sum of the masses of its constituent charges.*

"The nucleus of a hydrogen atom consists of one single free proton; its planetary electron is too far away to cause any effect, so that it is clear that we shall find the mass associated with 1 atom of hydrogen greater than one-fourth the mass of a helium atom or one-sixteenth the mass of an oxygen atom. The mass lost when 4 free protons and 2 free electrons are packed together to form a helium nucleus is roughly 0.7 per cent of the whole and it can be calculated that, if we take the value of the diameter of the electron given above, the protons must approach nearer than half of this to give so large a reduction. This means that the charges must be so closely packed that the electrons are actually deformed.

"The whole-number rule may now be simply translated into a statement that the mean packing effect in all atoms is approximately constant, and the unit of mass 1, when O equals 16, will be (mass of a packed proton) $+ \frac{1}{2}$(mass of free electron) $+ \frac{1}{2}$(mass of packed electron). The whole-number rule is not, and never was supposed to be, mathematically exact, for this would imply an identical packing effect in the case of all atoms, an exceedingly improbable supposition. It is almost certain that atoms of some elements, such as nitrogen, weigh slightly more than a whole number (looser packing), while those of iodine or caesium, for example, may weigh slightly less (closer packing). The limit of accuracy so far attained in mass-spectrum measurements is not sufficient to detect a change of the order expected, except in the case of hydrogen, where the variation in mass is exceptionally high."

25. The Arrangement of the External Electrons of an Atom.— Before we undertake the discussion of this topic, let us summarize the facts we have learned concerning the structure of atoms.

(1) Atoms are constructed of unit positive charges of electricity (protons) associated with unit charges of negative electricity (electrons).

The foregoing statements give us the key to chemical combination. When the atoms combine to form molecules they must do so either by the direct transfer of electrons from one atom to another or by the sharing of electrons with one another. The valence of an atom signifies the number of electrons that may be gained, lost, or shared when the atom enters into chemical union. Some atoms exhibit a strong tendency to lose electrons, others to gain electrons, while still others may either gain or lose a certain number of electrons. When atoms of the same type unite to form molecules, they usually do so by the sharing process.

Atoms like sodium and potassium are said to be strongly positive because they readily give up electrons and show little tendency to gain them. In like manner, atoms like chlorine and bromine are said to be strongly negative because they readily gain electrons and exhibit little tendency to lose them.

When 2 or more atoms have united to form molecules, the oppositely charged particles are held together by electrostatic attraction. Hence, we get Na^+Cl^-, Na^+Br^-, K^+Br^-, and K^+Cl^- when these strongly charged atoms are brought together under suitable conditions. With this in mind, it is easy to see why sodium and chlorine cannot be used interchangeably in compounds, since they do not have the same valence characteristics although the valence of each is 1. A compound of the type just mentioned is strongly "polar"; that is, it has a positive end and a negative end, like a magnet, and if such a compound is electrolyzed the two parts are attracted to opposite poles. A strongly polar compound does not usually exist in individual molecules in concentrated solutions, but tends to build itself into aggregates or into structures characteristic of crystalline solids.

Carbon and nitrogen are good examples of elements that can either lose or gain electrons with almost equal ease. The outer shell of the carbon atom lacks 4 electrons of a complete octet. In this case, the atom may either give up the 4 electrons of the outer shell or acquire 4 extra electrons from some other atom, thereby completing its octet. For example, a carbon atom can give up its 4 electrons to each of 2 oxygen atoms, thereby completing their octets and forming a molecule of carbon dioxide; in this case carbon has a valence of $+ 4$. Or, it can take up an electron from each of 4 hydrogen atoms, and in so doing complete its own octet; in this instance carbon has a valence of -4. Nitrogen has 5 electrons in its incomplete outer shell. It can

reach a stable condition either by gaining 3 electrons, thereby bringing its outer shell up to a stable 8, or, by losing 5 electrons, it can acquire the stable condition of helium. Hence, we find the most stable inorganic compounds of nitrogen to be the ammonium salts, in which it exhibits a negative valence of 3, and the nitrates, in which it has a positive valence of 5.

One of the most interesting features of the octet theory is the manner in which it explains the valence that is represented by the union between atoms of *non-polar valence*. There is a large group of compounds that seem to have been formed without the atoms transferring any of their valence electrons from one to another. In this type of union the valence electrons are assumed to be *shared* between two neighboring atoms in such a way as to hold the two atoms together. Organic compounds are very largely of this type. Molecules that are made up of atoms of the same kind are also non-polar.

Langmuir has developed an equation concerning the octet theory of valence, which enables one to determine the structure of non-polar compounds:

"By sharing a single pair of electrons 2 octets can be formed, since 2 octets holding a pair in common require only 14 electrons. This is clear if we consider the atoms as 2 cubes with electrons at each of the 8 corners. When the cubes are placed so that an edge of one is in contact with an edge of the other, a single pair of electrons at the ends of the common edge will take the place of 4 electrons in the original cubes. For each pair of electrons held in common between 2 octets there is a decrease of 2 in the total number of electrons needed to form the octet.

"Let e represent the number of electrons in the outside shell of the atoms that combine to form a molecule. Let n be the number of octets that are formed from these e electrons, and let p be the number of pairs of electrons which the octets share with one another. Since every pair of electrons thus shared reduces by 2 the number of electrons required to form the molecule, it follows that

$$e = 8n - 2p$$

or

$$p = \frac{8n - e}{2}.$$

"This simple equation tells us in each case how many pairs of

electrons or chemical bonds must exist in any given molecule *between the octets formed*. Hydrogen nuclei, however, may attach themselves to pairs of electrons in the octets which are not already shared. For example, in the formation of hydrogen fluoride from a hydrogen atom and a fluorine atom there are 8 electrons in the shell ($e = 8$). We place $n = 1$ in the above equation and find $p = 0$. In other words, the fluorine atoms do not share electrons with each other. The hydrogen nucleus, having given up its electron to the fluorine atom, attaches itself to one of the pairs of electrons of the fluorine octet, and thus forms a molecule having a relatively weak external field of force. As a result hydrogen fluoride is a liquid of low boiling-point instead of being salt-like in character."

The equation given above is applicable to all types of compounds. For example, the oxygen atom contains 8 electrons outside the nucleus; 6 of these are in the second shell. It thus needs 2 electrons to complete its octet. When oxygen atoms unite with each other to form molecules, O_2, they complete their octets by sharing electrons. By using the equation developed above, we can tell how many pairs of electrons are thus shared, and from these data we are able to determine the structure of the oxygen molecule. The equation was given as $p = \dfrac{8n - e}{2}$. In the case of oxygen, $n = 2$ and $e = 12$. Therefore,

$p = \dfrac{16 - 12}{2}$, or 2, which means that 2 pairs of electrons are held in common by the two octets. In other words, each octet shares 2 of its electrons with the other octet, both octets thereby being made complete. The external electrons of the oxygen molecule may therefore be represented by the following diagram:
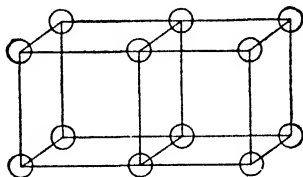
FIG. 5.—A Molecule of Oxygen.

27. " Isosteric " Compounds.—An unexpected and yet very interesting discovery was made by Langmuir while he was working out the structure of different compounds by the use of his octet theory. He

found that certain cases seemed to show identical configuration and number of electrons, and from this he reasoned that these substances should have similar properties. Upon looking up the known properties of the compounds in question he found his theory confirmed to a remarkable degree. For example, carbon dioxide and nitrous oxide, two gases of dissimilar chemical composition and properties, are found to possess extraordinary similarity in their physical properties. A molecule of each gas contains 16 electrons arranged around 3 atomic nuclei, as shown in Fig. 6.
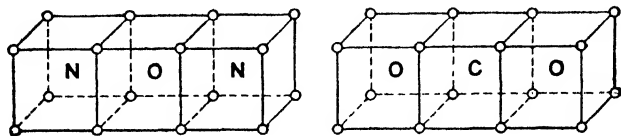
FIG. 6

Observe that the central atom in each molecule shares 4 pairs of electrons, 2 pairs being shared with each of the adjacent octets. The following table shows how nearly identical are the physical properties of these two gases:

TABLE V

PHYSICAL CONSTANTS OF N_2O AND CO_2

Constant	N_2O	CO_2
Boiling-point, abs.	183.2°	194.7°
Critical temperature, abs.	309.5°	304°
Critical pressure, atmos.	71.65	72.85
Critical density	0.454	0.448
Density of liquid, −20°	0.996	1.031
Density of liquid, +10°	0.856	0.858
Solubility in water, 5°	0.205	0.277
Viscosity, 0°, $\eta \times 10^7$	1408.	1414.
Refractive index of liquid	1.193	1.190
Dielectric constant of liquid, 0°	1.598	1.582
Magnetic susceptibility of gas at 40 atmos. $\times 10^{-6}$, 16°	0.12	0.12

N_2 and CO constitute another pair of gases which show a similar relationship.

Langmuir has termed this type of substances *isosteric* substances.

He considers the discovery of isosterism as another proof of the correctness of the octet theory in fundamental principle.

For further details of the octet theory the papers of Langmuir should be consulted.

28. Modern Conception of the Term " Element."—The exact idea conveyed by the term "element" in chemistry and physics has given rise to no end of trouble to scientists for countless generations. Dalton postulated that: An element is a substance that is made up of atoms similar to one another in weight and properties, and that cannot be resolved into further components by any known chemical process. On the basis of this postulate it has been possible to work out the relative atomic weights of the atoms of different elements—the system of atomic weights in use up to date. Nevertheless, modern research has revealed the fact that different atoms of a given element are often of different weights. In other words, the atomic weight of an element is now known to represent merely the *average weight*, relative to O = 16, of the several different types of atoms. It is a constant quantity simply because these different types of atoms, as they are found in nature, are always associated in a constant proportion, and because their chemical properties are so nearly alike that we have no chemical means of separating them.

The idea that the atoms of an element might be heterogeneous in respect to their relative weights, while only recently proved correct by Aston and others, was conceived many years ago by other chemists. Sir William Crookes, in 1886, made the following statement:[22]

"I conceive, therefore, that when we say the atomic weight of, for instance, calcium is 40, we really express the fact that, while the majority of calcium atoms have an actual weight of 40, there are not a few which are represented by 39 or 41, a less number by 38 or 42, and so on. . . . This may seem an audacious speculation, but I do not think it beyond the power of chemistry to test its feasibility."

The extraordinary insight that Sir William Crookes had into the nature of things was borne out by subsequent discoveries. The discovery of the electron as a constituent of the atom, ten years later, gave the first real proof that the atom is not the ultimate particle of matter. The discovery of radioactive elements and associated discoveries proved that some atoms are continually undergoing change. All of these facts proved that an atom is a very intricate structure, and

[22] Nat., **34**, p. 423, 1886.

led up to the present-day electrical theory of matter. The final proof of Crookes' prognostication was obtained when Sir J. J. Thomson and F. W. Aston developed the "positive-ray" method of analysis. By this means they proved that many non-radioactive elements, such as chlorine, mercury, krypton, magnesium, zinc, cadmium, calcium, tin, and a number of other elements are each a mixture of two or more isotopes.

The discovery of isotopes brings us face to face with two possible alternatives regarding what we should consider as an element. The first is to call each isotope, as it is discovered, a new element. This solution does not seem practicable, however, for all of the isotopes of any given element, such as chlorine, are, so far as our present-day knowledge goes, identical in their chemical properties, occupy the same position in the periodic table, and have the same atomic number. In other words, the atoms of a given element may differ in weight, but are believed to be identical in all their chemical properties. The second and more practicable solution of the difficulty is to associate the term element with the atomic number, i.e., to consider the term element to mean a substance which has definite chemical and spectroscopic properties, which may or may not be a mixture of isotopes, but which has a definite atomic number. On this basis there would be, corresponding to Moseley's numbers, 92 possible elements, of which 87 are known.

29. Mass Unchanged in Chemical Reactions.—It is very important to remember that one of the fundamental laws of chemical science is that of the conservation of mass. When a chemical reaction takes place, the properties of the substances produced are usually different from those of the reactants: a solid may become a liquid or a gas; a liquid may become a gas or a solid; a poisonous gas, like chlorine, may combine with a strongly corrosive metal like sodium and form a highly nutritious substance (sodium chloride); and so on throughout the entire list of properties, all of which may change except mass.

The question as to whether there is any change in weight in chemical reactions has been thoroughly investigated by Landolt.[23] His experiments proved that the weight of the products of a chemical reaction is exactly equal to the weight of the constituents before the reaction.

The conservation of mass, then, is the one unchangeable property that matter exhibits when undergoing chemical transformation. It is

[23] Zeit. phys. Chem., **12**, 1, 1893; **55**, 589, 1906; **64**, 581, 1908.

on this principle that the science of quantitative chemical analysis is based.

An important corollary of this law is *that the total amount of matter in the universe remains constant throughout all time.* In accepting this statement, it is well to keep in mind the modern conception of the atom. In other words, when we speak of the total quantity of matter in the universe remaining constant, we mean that the total quantity of positive and negative electricity (protons and electrons) is always constant.

30. Laws of Chemical Combination.—The enunciation of the law of conservation of mass and the introduction of the balance into chemical procedures resulted in the establishment of several laws of chemical combination. The first of these laws, which is known as the law of constant proportions, may be expressed as follows:

Any given chemical compound always contains the same constituents, and there is a constant proportion between the masses of the constituents present.

The validity of this law was called in question by Berthollet, who, from a series of experiments that he had performed, had become convinced that chemical reactions are largely controlled by the relative amounts of reacting substances; hence he contended that when a chemical compound is formed by the combination of two elements, for example, the proportion of one of the elements in the compound will be the greater the more of that element there is available. This led to the celebrated controversy between Berthollet and Proust (1799–1807), which ended in the firm establishment of the law of constant proportions. All subsequent investigation has strengthened our faith in this law.

In certain cases, it has been found that elements may combine in more than one proportion to form true chemical compounds. For example, Dalton found by analysis that two compounds of carbon and hydrogen—methane and ethylene—contain the elements in the ratios 3 : 1 and 6 : 1. The latter compound evidently contains twice as much carbon (by weight) with respect to hydrogen as the former.

A large number of compounds were examined, and similar simple ratios between the masses of the constituent elements were found. From these and similar facts Dalton arrived at the law of multiple proportions, which may be stated thus: *When two elements unite in*

more than one proportion, for a definite amount of one element there is a simple ratio between the amounts of the other element.

There is a third law of chemical combination, which includes the other two. It has been found possible to assign to every element a definite relative weight with which it enters into chemical combination. In other words, if we determine the weights of different substances which combine with any given weight of a definite substance, these weights, or simple multiples of them, represent the quantities of the different substances which will combine with one another. The weights of the elements which combine with one another are termed their *combining weights*, and hence the law is known as the law of combining weights. It is usually stated as follows: *Elements combine in the ratio of their combining weights, or in simple multiples of this ratio.* The combining weights are found by direct analysis. Water, for example, contains 8 parts, by weight, of oxygen to 1.008 of hydrogen; hydrogen sulphide contains 16 parts, by weight, of sulphur to 1.008 of hydrogen; and hydrogen chloride contains 35.46 parts, by weight, of chlorine to 1.008 of hydrogen. It is evident from these three analyses that 16 parts of sulphur and 35.46 parts of chlorine are equivalent to 8 parts of oxygen, and that these numbers represent the ratios in which these elements may be substituted for each other in chemical compounds. If we were to adopt arbitrarily the combining weight of hydrogen as unity, the figures for the other elements in question would be slightly different from those given above, but the ratio would be the same. It is clear, therefore, that we could adopt any element and any weight as the standard. Chemists have decided to adopt 8 grams of oxygen as the standard of comparison. With this value as our standard, a system of chemical equivalents has been built up. In silver oxide 8 grams of oxygen are found to be combined with 107.88 grams of silver; the latter value is therefore the combining weight of silver. It was stated above that hydrogen chloride contains 1.008 grams of hydrogen and 35.46 grams of chlorine. Now, silver chloride contains 107.88 grams of silver combined with 35.46 grams of chlorine. In sodium chloride we find 23 grams of sodium combined with 35.46 grams of chlorine. Finally, in sodium oxide we find 23 grams of sodium combined with 8 grams of oxygen. And so we might go on indefinitely.

The combining weights are often termed *chemical equivalents*, and since 1.008 grams of hydrogen combine with 8 grams of oxygen, it is customary to consider the chemical equivalent of an element as that

quantity of it which combines with, or displaces, 1.008 parts, by weight, of hydrogen.

31. Atomic Weights.—It is evident from the preceding that the atoms of each given element have a definite and invariable average weight; and that molecules are formed by the combination of definite whole numbers of atoms. Furthermore, the relative weights in which elements combine with each other are determined by the relative weights of their individual atoms.

In determining atomic weights we must choose some element as the standard. We have already seen that, in the estimation of equivalent weights of elements, hydrogen as unity was considered as the standard, only to be abandoned in favor of oxygen as 8. The reason for this change was that hydrogen does not combine directly with many of the elements to form stable compounds which can be analyzed, whereas oxygen, on the other hand, does combine with a large number of the elements, forming some of the most stable compounds with which we are acquainted. For the same reason, chemists have agreed to adopt the number 16 as the atomic weight of oxygen, and to use this as the basis of atomic weight estimations. By analysis we have found that each molecule of water contains 8 parts of oxygen to 1.008 parts of hydrogen; but the molecule of water is known to consist of two atoms of hydrogen combined with one of oxygen; hence the amount of hydrogen and oxygen in a molecule of water must be 2.016 parts of hydrogen to 16 parts of oxygen.

The usual method of determining the combining weight of an element, in terms of oxygen as the standard, is to determine the weight of the element that combines with a known weight of oxygen. The combining weight of the element is then calculated by the following simple proportion: Weight of O : weight of element : : atomic weight of O : combining weight of element. We have then to determine by other procedures how many atoms of the elements in question combine with one atom of oxygen. Most of the atomic weights have been determined in this way, but other methods have been used for verification.

As we have already learned, the "positive-ray method" of Thomson and Aston is by far the most accurate one that has yet been devised for estimating the true atomic mass of elements.

32. Chemical Formula, Weights, Equivalent Weights, and Gram-molecular Weights.—A *chemical formula* is a group of symbols showing

what kinds of *atoms* and how many of each are present in any given molecule. In general, the symbols of an element represent either the individual atom or as many grams of the element as there are units in its atomic weight. Such a weight is called a gram-atomic weight, or a gram-atom. Thus a gram-atom of sulphur is 32.064 grams. Hence, each formula shows the kinds of atoms and the number of each kind that are contained in a molecule of the given substance, and if we apply the gram-atom value to each symbol, then the whole formula represents a weight in grams of the substance that is equal to the sum of the atomic weights each multiplied by its subscript. This weight is called the *gram-formula weight*. Thus, the formula H_2SO_4 denotes $1.008 \times 2 + 32.06 + 16 \times 4 = 98.08$ grams of sulphuric acid. HCl denotes $1.008 + 35.46 = 36.46$ grams of hydrochloric acid; and Na_3PO_4 denotes $23 \times 3 + 31.027 + 16 \times 4 = 164.027$.

CHAPTER III

GENERAL PROPERTIES OF MATTER

1. States of Aggregation.—Matter occurs in different states or conditions known as *states of aggregation*. The three principal states of aggregation are the *gaseous state*, the *liquid state*, and the *crystalline state*. There is another important modification of the latter two states, known as the *colloidal state*. In addition, there are *amorphous solids, solid liquids*, and *liquid crystals*.

The three principal states—gaseous, liquid, and crystalline—differ from one another in many respects; but the most important difference is in the relative ease with which the particles can move among one another. If a substance is in the gaseous state there is comparatively little resistance to the movements of the molecules; the friction of one particle against another is relatively slight. In a liquid the internal friction is much greater than in a gas, hence the particles do not have the same freedom of movement that they do in a gas. Finally, if matter is in the crystalline state of aggregation the particles are relatively fixed, that is to say, crystalline bodies have a definite shape, and this persists until it is changed by some outside influence, such as subjecting the crystals to very great pressures.

2. Properties of Gases: The Gas Laws.—The gaseous form of matter is characterized by its tendency to fill completely and to a uniform density any available space. This tendency of a gas to fill the entire space at its disposal makes it possible to change the volume of the gas at will. This can be effected by simply changing the pressure to which the gas is subjected. For example, if a given quantity of a gas is placed within a cylinder fitted with a gas-tight movable piston, any increase in pressure (provided the temperature is kept constant) will cause the piston to descend and confine the gas within a proportionately smaller volume. It is evident that with increase of pressure the density of the gas becomes greater, and the inner friction of the molecules also increases.

From the foregoing remarks, it is evident that the quantitative laws which express the behavior of gases when external conditions of temperature and pressure are modified are largely independent of the nature of the gas. In consequence of this, general laws expressing the behavior of gases under varying conditions have been evolved. Owing to the fact that gases are much less dense than liquids and solids, the laws governing this state of aggregation are relatively simple.

(a) *Boyle's Law.*—In 1662, Robert Boyle enunciated the well-known law that at constant temperature the volume of a given mass of any gas is inversely proportional to the pressure.

If v represents the volume of any weight of a gas at the pressure p, and v_1 the volume of the same amount of the gas at the pressure p_1, then

$$pv = p_1v_1 = \text{constant for constant temperature}$$

(b) *Gay-Lussac's Law.*—If a gas is kept under constant pressure and its temperature is raised, its volume will increase. If the volume is kept constant as the temperature rises, the pressure will increase. Furthermore, the same rise of temperature has been found to produce in all gases the same increase in volume. Gay-Lussac, in 1801, discovered that all gases increase about $\frac{1}{273}$ of their volume at 0° for every rise of one degree temperature as expressed on the Centigrade system. This law may be formulated thus: *At constant pressure, the volume of a gas is directly proportional to its absolute temperature* (273 + temp. Centigrade above zero).

If the temperature were decreased continuously from 0°, we should finally reach a point at which the volume would be equal to zero. Provided the law of Gay-Lussac holds for such a low temperature, this point will be 273 Centigrade degrees below 0° C. (Long before this temperature could be reached, however, the gas would have passed into the liquid state and hence the gas laws would have ceased to be operative.) This point is called the *absolute zero*, and the temperature measured from it in Centigrade degrees is called the *absolute temperature.* This is usually designated by the letter T, the temperature according to the Centigrade scale being represented by t. Hence the absolute temperature $T = t + 273$, or $t = T - 273$.

All gases increase about $\frac{1}{273}$ (= 0.00367) of their volume at 0° C. for every rise of one degree in temperature. Now, if we keep the volume

constant and heat the gas to t degrees, the pressure p, at this temperature is calculated from the pressure, p_0, at $0°$ as follows:

$$p = p_0(1 + 0.00367t).$$

If the pressure is kept constant and the volume is permitted to increase through a rise in temperature, the volume at $t°$, V can be calculated from the volume at $0°$, v_0, as follows:

$$V = v_0(1 + 0.00367t).$$

If both the volume and pressure are permitted to change when the gas is warmed, the volume and pressure at, $t°$ p and v can be calculated from the pressure and volume at $0°$, thus:

$$pv = p_0v_0(1 + 0.00367t),$$

from which

$$v_0 = \frac{pv}{p_0(1 + 0.00367t)}.$$

This is the equation that is generally used for reducing a gas to normal conditions. If the volume, V, of a gas is observed at a given pressure, p, and temperature, t, we can calculate the volume, v_0, at $0°$ C. and normal pressure p_0, which is taken as 760 mm. of mercury.

The temperature coefficient of a gas—0.00367—can be determined either by keeping the pressure constant and measuring the increase in volume with rise in temperature, or by keeping the volume constant and measuring the pressure as the temperature rises.

If we represent the temperature as measured from the absolute temperature by T, the combined expression of the laws of Boyle and Gay-Lussac is

$$pv = \frac{p_0v_0}{273}T$$

Now $\frac{p_0v_0}{273}$ is usually represented by R, hence we have $pv = RT$.

(c) *Avogadro's Law.*—The law of Avogadro is usually stated thus: *All gases under the same conditions of temperature and pressure contain in unit volume the same number of molecules.* This is one of the most

important laws in chemistry, for by its aid we are able to determine the molecular weight of gases, or substances that can easily be converted into the gaseous form, from their densities. Since equal volumes contain the same number of molecules, the weight of these volumes must vary according to the weight of the molecules themselves. Another way of stating this law is as follows: *Under the same conditions of temperature and pressure, the gram-molecular weights of all gases will occupy the same volume.* The temperature of 0° C. and normal atmospheric pressure (760 mm. of mercury) is the standard that is usually chosen for comparative results. This is represenred as *N. T. P.*, and under these conditions the *gram-molecular volume* of all gases is 22.4 liters.

From Boyle's law it is evident that, if the volume is reduced to 1 liter at 0° the pressure will be 22.4 atmospheres.

The three laws of gas pressure can be combined in one expression, just as the laws of Boyle and Gay-Lussac are combined in the equation $PV = RT$.

We have stated above that the pressure exerted by a gram-molecule of a gas is 22.4 atmospheres when the gas occupies the space of 1 liter at 0°. If we substitute 22.4 for p_0v_0 in the equation $pv = \dfrac{p_0v_0}{273}T$, we get $pv = 0.082T$.

In other words, in the combined expression of the laws of Boyle, Gay-Lussac and Avogadro, R is equal to 0.082 (in liter atmospheres).

(*d*) *Deviations from the Gas Laws.*—Careful experiments by Amagat[1] and others on the behavior of gases at various temperatures and pressures have shown that the gas laws do not hold under all conditions. Deviations from Boyle's law were especially noticeable at high pressures; the change in volume being less than it should have been if the law held under these conditions. Amagat concluded from his experiments that the product of pressure and volume, *pv*, increases with increase of pressure for very high pressures. He experimented with hydrogen, nitrogen, carbon dioxide, oxygen, ethylene, etc., and found that *pv* increases continuously with the pressure in the case of hydrogen; but that for certain other gases, such as nitrogen and carbon dioxide, it first decreases, reaches a minimum, and beyond that point increases with increase of pressure. Amagat also found that the

---

[1] Ann Chim. phys. 5, 19, p. 345, 1880.

coefficient of expansion varies considerably from one gas to another, and varies for the same gas under different temperatures and pressures. In the case of ethylene, for instance, the coefficient of expansion increases with the temperature up to a certain point and then diminishes. At higher temperatures the coefficient becomes still less. This occurs with changes in temperature when the pressure is kept constant.

The laws of Boyle and Gay-Lussac are to be regarded, therefore, as approximations, which hold rigidly only under special conditions.

It was thought for a time that Avogadro's law did not hold for all substances under all conditions, but further research showed that when all the facts were correlated the exceptions were only apparent. The case of ammonium chloride will illustrate the point. The density of the vapors of this gas were found to be different from what would be expected from the law. It was found, however, that this gas ionizes at high temperatures, hence the number of particles was greater at such temperatures, and so naturally the law could not hold under these conditions.

It may be said, in general, that the laws are the more nearly obeyed the higher the temperature and the smaller the pressure, and, as regards the nature of the gas, the farther it is removed from the temperature of liquefaction. An *ideal gas* would be one that followed the gas laws in every way.

The general behavior of gases and the deviations from the simple laws may be satisfactorily accounted for on the basis of the kinetic theory of gases, which we shall now consider.

3. Kinetic Theory of Gases.—The first attempt to explain the properties of gases on the basis of the laws of mechanics, or "dynamics," was made by Bernoulli in 1758. Our present views of the subject, however, are due to the combined researches of Kroenig, Clausius, Maxwell, and Boltzmann. Of course, many others have either directly or indirectly added to the theory, which is known as the *kinetic theory of gases*. According to this theory, gases are made up of small, perfectly elastic particles (molecules) which are continuously moving in all directions in straight lines, the velocity being very great, and each particle moving independently of all others. The space actually occupied by the gas particles is considered to be very small as compared with the volume filled by them under ordinary conditions, thereby permitting the particles to move about free from one another's influence except when they collide. Owing to the rela-

tively large free space in which the particles move, each particle can usually travel a considerable distance before it meets with another particle. But whenever two particles do happen to collide, they rebound without loss of energy. The average distance through which a particle travels before colliding with another particle is known as its *mean free path*. These particles frequently strike, not only one another, but also the walls of the containing vessel, and thereby produce the so-called gas pressure. In other words, the pressure of the gas on the walls of the confining vessel is due to the blows or impacts of the gas particles against these walls.

4. Deviations from the Gas Laws Explained by the Kinetic Theory: Van der Waals' Equation.—We have seen that the gas equation ($pv = RT$) does not describe the behavior of any but "ideal" gases with respect to changes in volume with variations of temperature and pressure. The same gas does not behave in the same way at high and low pressures. Let us now examine the behavior of gases in the light of the kinetic theory and see whether any explanation of the exceptions to the gas laws can be found. According to the kinetic theory, gas pressure is due to the impacts of the molecules against the walls of the container. Now, the space in which these molecules move is evidently not the whole volume of the gas, but is equal to the volume of the gas minus the space occupied by the molecules. When the pressure is small, that is, when the volume of the gas is large, there are relatively few molecules in a large space, so that the space occupied by the molecules themselves is relatively so small that it is negligible. If, however, the volume is greatly diminished, the volume occupied by the non-compressible molecules will be relatively great. It is evident then that we must take into account the space occupied by the molecules themselves. If this volume correction is denoted by b, then the gas equation becomes

$$p(v - b) = RT$$

where b is a constant for each gas.

Another flaw in the equation $pv = RT$ is that no account is taken of any attractive force between the molecules. As a gas approaches liquefaction, the intermolecular attraction, which is almost absent in dilute gases, will certainly increase and act in the same sense as external pressure. Van der Waals has taken this into account, and has modified the gas equation accordingly. He has shown that this

correction is inversely proportional to the square of the volume, and since it augments the applied pressure, the expression, $p + \dfrac{a}{v^2}$ should be substituted for p in the gas equation, a being the constant of molecular attraction. The equation then becomes

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT.$$

This equation, which is known as *Van der Waals' equation*, explains many of the exceptions to the laws of Boyle and Gay-Lussac exhibited by gases. If the pressure is slight, $\dfrac{a}{v^2}$ becomes negligible, since the value of v is great, and the space occupied by the molecules b is likewise small. However, if a has a large value, the volume is appreciably diminished, and pv decreases, so that finally the influence of b begins to manifest itself.

5. Densities and Molecular Weights of Gases.—The determination of the relative densities of gases consists in determining the relative weights of equal volumes of gases at the same temperature and pressure. Since equal volumes of gases under the same conditions contain the same number of molecules, the densities stand in the same relation as the molecular weights. Thus, by means of Avogadro's law we can determine the molecular weights of substances in the gaseous state.

Some substance must be taken as the standard in determinations of the densities of gases. Air is usually chosen for this purpose, and the weights of equal volumes of other gases, at the same temperature and pressure, are compared with that of air. Hydrogen has also been used as a standard. Many prefer it to air, as it is more constant in its properties. The density of air is about 14.37 times the density of hydrogen; now, since the molecular weight of hydrogen is 2.016, we must multiply the density, referred to air as the standard, by 28.97, to obtain the molecular weight of the gas. If we represent the molecular weight of the gas by m, and the density, referred to air as the standard, by d, then:

$$m = d \times 28.97.$$

In this way the molecular weight of gases can be calculated from their densities.

6. General Characterictics of Liquids.—The most marked characteristic of the liquid state of aggregation is that a given mass of a liquid has a definite volume but no definite form. The volume of a liquid is dependent upon temperature and pressure, but to a lesser degree than is the volume of a gas. In other words, a liquid has its own definite volume, and this volume is only slightly changed by a change in conditions; it is slightly diminished by increase in pressure and slightly increased by increase in temperature; but the change in either case is small. This is undoubtedly due to the fact that in liquids the molecules are much less mobile than in gases, the distance between contiguous molecules being much less in liquids. One cubic centimeter of liquid water at 100° C., when converted into gaseous water, at the same temperature, occupies a volume of about 1700 c.c. This shows that, while the particles of a liquid move with comparative freedom about one another, the resistance to movement is enormous as compared with that found in gases. The *mean free path* of the molecules is much less in the liquid state. Furthermore, the very close proximity of the molecules to one another makes the effect of their mutual attractions and repulsions very pronounced and gives rise to the phenomenon of "surface tension." The molecules inside the liquid are attracted in every possible direction by their neighbors, while those at the surface are attracted only inward and towards each other. This makes the surface layer relatively dense and difficult to penetrate, just as if an elastic skin were drawn all about the liquid. The surface tension always acts so as to make the total surface of the liquid as small as possible.

7. Connection between the Gaseous and Liquid States.—If at constant temperature a gas is subjected to a constantly increasing pressure, its state may change in one of two ways, according to conditions:

(*a*) The volume at first diminishes more rapidly than the pressure increases, next in the same ratio, and finally more slowly. When the pressure becomes very high a further increase has very little effect upon the volume.

(*b*) In this case the volume changes more rapidly than the pressure up to a certain point, then at this pressure the gas ceases to be homogeneous, and partial liquefaction results. For a constant temperature, the pressure at which liquefaction occurs is constant for a given gas, while the volume decreases steadily until liquefaction is complete.

While the process of liquefaction is going on, it is impossible to increase the pressure. Then the law of compressibility of liquids applies, i.e., liquids can be slightly compressed, but great pressure is required.

The temperature is undoubtedly the determining factor that causes a gas under compression to follow *a* or *b*. There is a certain temperature, which depends upon the nature of the gas, above which process *a* will be followed; and below this point, process *b*. That a gas may behave in either of the above ways was first definitely recognized by Andrews in 1869, in connection with his experiments on the liquefaction of carbon dioxide. If CO_2 gas is compressed, the temperature being kept at 0°, the volume changes more rapidly than the pressure until a pressure of 35.4 atmospheres is reached; at this pressure liquefaction occurs. As the temperature is raised, it is found that a higher pressure is required to liquefy the gas, until at temperatures of 30-32° C., or higher, it is impossible to condense the gas to the liquid state. The temperature above which it is no longer possible to liquefy the gas, Andrews termed the *critical temperature*. In like manner, the pressure required to liquefy the gas at the critical temperature, he termed the *critical pressure*; and the volume occupied by the gas or the liquid under these conditions, he called the *critical volume*.

8. Boiling-point of Liquids.—When a liquid is heated under a definite external pressure (for example, water under the pressure of one atmosphere), its vapor pressure increases with rise in temperature until finally a temperature is reached at which bubbles of vapor rise from within the body of the liquid and escape into the space above it. The liquid is then said to boil. The *boiling-point* of a liquid may be defined as the temperature at which its vapor pressure becomes equal to the external pressure. The temperature indicated by a thermometer placed in a boiling liquid is generally not its true boiling-point. Very special precautions must be taken in order that a liquid may boil under equilibrium conditions. In order to secure equilibrium and measure the temperature corresponding to it, an intimate mixture of vapor and liquid under the same total pressure must be brought into contact with the thermometer, and this is a difficult thing to do. Hence, in determining the boiling-point of a liquid, it is customary to place the thermometer in the vapor just above the liquid, rather than in the liquid, for the latter is more likely to become superheated.

If the external pressure is reduced by means of an air pump, the boiling-point is correspondingly lowered, i.e., the liquid will boil at a

lower temperature, just as it does at the top of a high mountain. This principle is made use of in evaporating many biological fluids, which would be injured if heated to the boiling-point at ordinary atmospheric pressure. The converse of this principle is applied in the *pressure cooker*, or *autoclave;* i.e., by increasing the pressure, the temperature at which the liquid boils is raised. There are also many biological applications of this principle. For instance, the autoclave is widely used in bacteriological laboratories.

9. Viscosity and Fluidity.—The internal friction of resistance exhibited by the molecules in moving around in the interior of a body is termed *viscosity*. The viscosity of gases is very low; that of crystals, very high; while that of liquids includes a wide range of variations. Some liquids, like ether, are very "mobile," while others, such as tar, are very *viscous*.

The "coefficient of viscosity" or, simply, the viscosity of a liquid can be measured by comparing the relative time which the given liquid requires to flow through a vertical capillary tube with the time of flow of a corresponding quantity of some standard liquid under the same experimental conditions.

The reciprocal of viscosity is called fluidity.

$$\text{Viscosity} = \frac{1}{\text{fluidity}},$$

or, what amounts to the same thing,

$$\text{Fluidity} = \frac{1}{\text{viscosity}}.$$

The fluidity is sometimes a more convenient quantity to measure than the viscosity. The principle of the method given above for the determination of the coefficient of viscosity is really the determination of the fluidity.

10.—General Characteristics of the Crystalline State.—The third state of aggregation is known as the crystalline state. We have already learned that when any gas is cooled below a certain point it passes over into the liquid state. In like manner, when a substance in the liquid state is gradually cooled, a point is usually reached where the crystalline state appears in the system. This state of aggregation is characterized by the fact that the bodies have a definite shape. Rigidity is

another characteristic. Work of some kind must be brought to bear upon the crystals in order to change their rigid, definite shape. A more critical examination shows that every pure substance in this state has a definite geometric form, which is one of the characteristic properties of the substance. The faces of crystals of the same substance may vary in size and shape; but if the crystals possess the same chemical composition, and are at the same temperature, the interfacial angles have the same numerical value. In other words, the angles between similar faces of crystals of the same substance are precisely the same, and are characteristic of that substance. It follows that the crystalline form of a substance is not determined by the absolute position nor by the sizes of the faces of the crystals, but rather by the dimensions of the interfacial angles.

It was mentioned above that most crystalline substances are characterized by having a definite shape that is so rigid as to require mechanical force to effect deformation. For this reason it has been customary in the past to speak of the crystalline state as the *solid* state of aggregation. The property of solidity, however, is not restricted to the crystalline state, but is possessed to a remarkable degree by some liquids, the so-called *amorphous solids*, which may be considered as liquids possessing great viscosity and small elasticity. The physical properties of amorphous solids have the same values in all directions, in contrast to crystals, which are usually different in different directions. When a crystalline substance is gradually heated under constant pressure (that of one atmosphere, usually) it either completely evaporates, thereby changing into a gas, or else it loses its crystalline form and changes into the liquid state as soon as a definite temperature is attained. This temperature, which depends slightly upon the pressure, is called the *melting-point* of the substance and is characteristic of it. On the other hand, when an amorphous solid is heated, it gradually softens and finally acquires the properties of a liquid, but there is no definite temperature at which the change from the solid to the liquid state takes place, as in the case of crystals.

11. Crystal Systems.—In a crystal the particles are arranged in a perfectly orderly manner, and fulfill the condition that the arrangement about any one point is the same as about any other point. Since the crystal is the geometric expression of the attraction of the particles that make up the substance, there are naturally many types of crystals. Nevertheless, it has been found that the countless different forms can

be considered as belonging to one or the other of six systems, which are as follows:

First, the *regular system*, having three axes, all of equal length and all making right angles with each other.

Second, the *tetragonal system*, having two axes of equal length and the third axis either longer or shorter, but all three making right angles to each other.

Third, the *hexagonal system*, having three axes of equal length crossing each other at angles of 60° and a fourth axis, either longer or shorter, at right angles to the others.

Fourth, the *orthorombic system*, having three axes of unequal length but all forming right angles with each other.

Fifth, the *monoclinic system*, having three axes of unequal length, two of which cross at right angles, while the third is at right angles to one but not to the other.

Sixth, the *triclinic system*, having three axes of unequal length but no right angles.

The above system, which was proposed by Weiss in 1809, enables one to recognize and describe practically any type of crystal that one may come across.

12. X-rays and Crystal Structure.—As was pointed out in Chapter II, Art 8, much light has recently been shed upon the question of the arrangement of the atoms and molecules within crystals, by the investigations that have been carried out on the transmission and reflection of X-rays by crystalline substances. The X-rays are now recognized to be of the same nature as ordinary light rays but to have extremely short wave-lengths. In considering the nature of X-rays, Professor Laue (Professor of Physics in the University of Zurich, Switzerland), reasoned that if the X-rays are similar to light rays, the successive rows of molecules in a crystal should behave toward these very short waves in the same way as a grating spectroscope does to ordinary light rays, that is, diffraction effects and X-ray spectra should be obtained. Experiment not only confirmed this conclusion, thereby giving the desired information concerning the X-rays themselves, but at the same time gave us pictures of the internal structure of crystals.

Several methods have been devised for the purpose of studying crystal structure by the use of X-rays. The basis upon which all of these methods are founded is the discovery of Laue (1912) that the

regular arrangement of the atoms in crystals represents a three-dimensional diffraction grating for the X-rays. Three of these methods will be briefly described.

In the Laue method, a beam of X-rays is passed through a thin section of a crystal and then allowed to impinge upon a photographic plate where a diffraction pattern of the crystal lattice is produced. After several measurements have been made, a model of the space arrangement of the atoms in the crystal can be worked out.

The second method we owe to W. H. Bragg, and his son, W. L. Bragg.[2] Their method is based upon the assumption that the diffraction of X-rays by the atoms of a crystal can be studied more easily by employing the crystal as a reflection grating, rather than as a transmission grating, as in the method of Laue. In order to apply this principle they devised an instrument known as an X-ray spectrometer. The crystal is mounted in this instrument so that it can be used as a *reflection* grating. It is so arranged that the various faces can be successively employed. The positions and intensities of the lines of the X-ray spectra obtained by reflection from the different faces of the crystal serve to determine the relative positions of the atoms within the crystal.

In the third method, which was devised by Hull, the crystalline substance is employed in the form of a powder. The main advantage of this method is that a very small amount of the substance is all that is required for the determinations, whereas both of the other methods require a crystal of sufficient size to be manipulated. In the determination a narrow beam of monochromatic X-rays is passed through a small quantity of the finely divided crystalline substance, when a characteristic spectra, made up of fine lines, will be produced on a photographic plate. These lines are due to the reflections of the X-rays from the different faces of the minute crystalline particles. From the positions and intensities of the lines, the crystal structure can be determined.

By means of the foregoing methods we have obtained much valuable knowledge concerning the structure of crystals; in fact, it has been possible to determine the exact structure in many cases. For example, the arrangement of the atoms in a crystal of rock-salt (NaCl) has been found to be similar to the model shown in Fig. 7a, in which the white circles represent sodium atoms and the dark circles chlorine

[2] X-Rays and Crystal Structure, London, 1924.

atoms. It has been found that the efficiency of an atom as a reflector of X-rays is proportional to its atomic weight, hence the relative positions of the atoms in the crystal are revealed by the relative intensities of the reflected rays. By applying this principle, the atoms of sodium and chlorine have been shown to alternate as shown in Fig. 7a. Furthermore, since the wave-length of the X-rays used in studying the crystals is known, the distance between the planes of the atoms can be calculated. For rock-salt this distance is found to be

$$d_{(100)} = a \qquad d_{(110)} = a/\sqrt{2} \qquad d_{(111)} = 2a/\sqrt{}$$
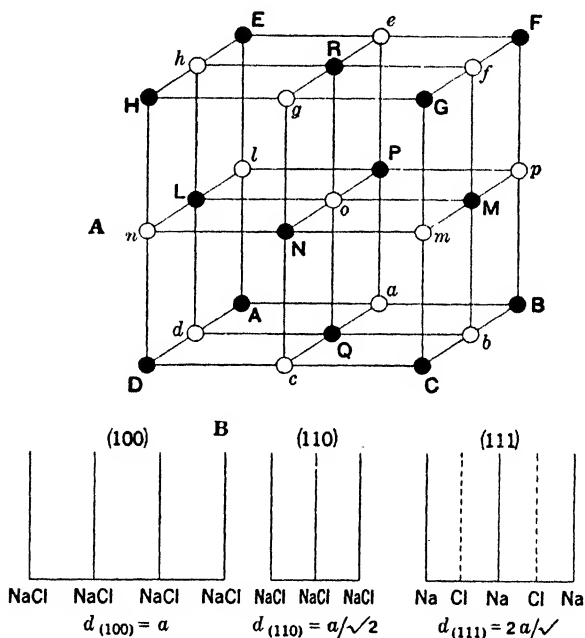
FIG. 7.

2.81×10^{-8}cm. Such an arrangement of the atoms in planes in rock-salt is shown diagrammatically in Fig. 7b.

The most remarkable part of all these experiments is that in the crystal the individuality of the molecule seems to have disappeared and the atom seems to have become the unit of structure. This is especially true of crystals of strongly polar substances, such as the halide salts. To be more exact, the unit of structure in the polar substances is the ion. For example, the cubical crystal of potassium chloride is made up, not of a succession of independent molecules of KCl, but of a rectangular lattice work of alternate K and Cl ions placed at

equal distances from each other. If we examine the model of NaCl (Fig. 7a), we shall see that each sodium ion in the crystal is surrounded by 6 equidistant chlorine ions placed at the corners of an octahedron. Each chlorine ion in the space lattice occupies a similar central position with respect to 6 equidistant sodium ions. In fact, we should really call the whole crystal a gigantic molecule.

There are other cases in which some of the units of crystal structure are groups of atoms, or parts of molecules. In other words, just as we have groups of atoms acting *en masse* in solutions of electrolytes, we have groups of atoms, not whole molecules, maintaining their identity in the crystal. Note the arrangement of the atoms in a crystal of calcite ($CaCO_3$) as depicted in Fig. 8. Each carbon atom has 3 oxygen atoms symmetrically disposed about it.

⊗   ●   ○
Ca   C   O

Fig. 8.

In the non-polar compounds the molecule seems to be the unit of crystal structure.
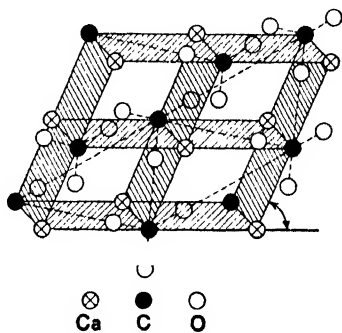
13. The Melting-point of a Crystalline Substance a Criterion of Purity.—Of all the methods of identifying a crystalline substance and testing its purity, no one is so frequently made use of by the chemist as the melting-point method. The temperature at which a substance melts is a characteristic constant for the substance, and this fact is used as a means of identification. Furthermore, if a substance is *not* pure it will not melt sharply at one temperature, hence this test can be used to determine the purity of the substance.

CHAPTER IV

ENERGY TRANSFORMATIONS IN LIVING MATTER

1. The Cell as an Energy Transformer.—If we apply the ordinary physical and chemical procedures to the study of living matter we are sure to observe that even the simplest forms of life are continuously undergoing change. No living thing is absolutely still. It may alter its position in relation to its environment; it may grow; it may undergo changes in internal structure. The changes in internal structure are of two distinct types. There is a building up of matter, and a breaking down of matter. If the building-up process predominates over the breaking-down process, the organism grows; if the two processes just balance, the organism exists; if the breaking-down process predomininates, the organism tends to disintegrate. Complete immobility betokens death. In other words, the most striking characteristic of living organisms is the perpetual state of change which they show.

Now, it is a rule in elementary mechanics that in order to effect changes, work must be done, and that the capacity to do work is due to the possession of something that is called energy. Some may argue that living matter is endowed with some peculiar " vital " force and that the vital processes do not necessarily obey the laws of physics, chemistry and mechanics, but we have, up to the present time, no evidence of such a " peculiar " force. To quote the great French physiologist, Claude Bernard, on the subject of vitalism: " There is in reality only one general physics, only one chemistry, and only one mechanics, in which all the phenomenal manifestations of nature are included, both those of living bodies and those of inanimate ones. In a word, all the phenomena which make their appearance in a living being obey the same laws as those outside it. So that we may say that all the manifestations of life are composed of phenomena borrowed from the outside cosmic world, so far as their nature is concerned, possessing, however, a special morphology, in the sense that

they are manifested under characteristic forms and by the aid of special physiological instruments."

" It must be remembered, however, that although living matter, after it is formed, is known to obey the laws of chemistry and physics, up to the present time, no physico-chemical system has been met with having the same properties as those known as vital; in other words, none has, as yet, been prepared of similar complexity and internal coördination."

Viewed, therefore, from the physico-chemical point of view, the living cell is a peculiarly constructed energy machine or energy transformer, through which a continual flux of energy ceaselessly goes on, and the entire life of the cell is an expression of variations and alterations in the rates of flow of energy, and changes in equilibrium or balance between the various types of energy.

2. General Concept of Energy.—Before attempting to discuss the source of the energy of living matter, or endeavoring to consider how it operates in vital phenomena, it may be well to consider briefly the general concept of energy and the laws that govern its manifestations.

(a) *Definition of Energy.*—Energy is usually defined as the power of doing work that a body possesses by reason of its state. In a broader sense, however, it may be considered as anything that produces an effect on our senses. For example, if we see an object, it is through the medium of radiant energy, light, which acts on the retina; if we hear, it is through the medium of kinetic energy acting on our eardrums; if we smell or taste, the sensation may be traced to chemical energy producing effects on the sense endings; if we feel, it is kinetic energy again, for we impart motion to our fingers, and feeling consists in compressions or displacements of nerve endings on our fingers.

(b) *Changes in Properties.*—Wood burns, and as a result of the combustion we get a totally different set of properties from that which we had originally. Water boils, and a different set of properties appear. Now, whenever we note a change in properties, what we actually observe is a transformation of energy.

For convenience, we may divide changes in properties into two classes, i.e., those that are only temporary and those that are relatively permanent. The former we call *physical changes*; and the latter, *chemical changes*. As an example of physical change we may take the changes that occur in a wire when electrical energy passes through it. While the current is passing through the wire it becomes possessed of

new properties. It becomes heated, and will deflect the needle of a compass, etc.; but as soon as the current ceases these properties no longer manifest themselves, and we have the wire in the original state. On the other hand, if we pass an electric current through a solution of the wire in acid and water, the electrical energy electrolyzes the mixture, and new substances appear at the electrodes. These new substances do not disappear when the current ceases. This type of change is permanent and is known as a *chemical change*.

(c) *Classification of Energies.*—Although we do not know how many forms of energy there are, for all practical purposes, we may consider all manifestations of energy under the following eight forms: kinetic, linear, surface, volume, heat, electrical, chemical, and radiant. Of course, other manifestations are possible, and we may use whatever classification suits our purpose best. For instance, for some purposes it is convenient to include the first four forms in this list under the one name, *mechanical*, and then to subdivide this into *kinetic* and *potential*, kinetic energy being considered as due to motion, and potential energy to position.

Energy is measured by its power to do work, hence the same unit of measurement is used, namely, the erg.

It is convenient at this point to include the following definitions: *Force* is whatever changes, or tends to change, the motion of a body by altering its direction or its magnitude. A force acting upon a body is measured by the momentum it produces in its own direction in unit time. The unit of force is the *dyne*, i.e., the force which, acting upon a gram for one second, would give it the velocity of one centimeter per second. Since a weight is used in this definition, only those forces are comparable that are measured at the same place. To avoid this difficulty, it is simply necessary to multiply the result by *grams*, the acceleration due to gravitation at that place. At Washington, a body falling freely for one second acquires the velocity of 980.10 cm. In other words, the weight of one gram at Washington is equal to 980.10 dynes, or one gram-centimeter is equal to 980.10 dynes.

Work.—The unit of work is that amount of work which is done when unit force is overcome through unit distance. This is called, in the absolute system, the *erg*.

We have, then,

$$dynes \times centimeters = ergs.$$

The maximum work is that amount which can be produced under ideal conditions.

We have no means of knowing what is the total amount of energy in any one system, for all our measurements are relative. The excess energy that any system contains in a given state over what it contains is some standard state is called the *available energy*. It is determined by allowing a body to go from the given state to the standard state in such a way that all the difference in energy appears in one form, for example, as heat. If we measure this heat, we have the available energy of the body in terms of heat energy.

(*d*) *The Factors of Energy.*—It has been found that the different forms of energy can each be expressed in the form of the product of two factors. One of these, called the "intensity factor," determines whether a transference of energy will occur or not, and in what direction; the other one, called the quantity or "capacity factor," determines to what extent the change will take place. Bodies having identical intensity factors are in equilibrium, hence no transfer of energy takes place between them. If two substances have different values for their intensity factor, then a transference of energy will take place until each has the same value and equilibrium is established. For example, suppose we have two vessels which are connected at the bottom with a tube having a pinchcock; if we place a liter of water at 50° C. in each vessel and then open the pinchcock, the energy of the system will be twice as great as that of the contents of one vessel, owing to the doubling of the capacity factor; the intensity factor (the temperature), on the other hand, is not altered. If, however, we were to place a liter of water at 50° C. in one vessel, and a liter of water at 100° C. in the other vessel, and were then to open the pinchcock, an exchange of energy (heat) would take place until the temperature in both vessels was the same, i.e., until their intensity factors were equal.

In electrical energy, the intensity factor is the difference of potential or electromotive force, while the capacity factor is the current.

(*e*) *Convertibility of Energy.*—The energy of a body may change from potential to kinetic, and conversely. Suppose work has been done on a weight of mass m gm., sufficient to lift it to a height h cm. against gravity. It then possesses potential energy equal to mgh ergs. If it is allowed to fall, it loses potential energy and gains energy of motion. After it has fallen a distance s, its velocity is given by the equation $v^2 = 2gs$. Its kinetic energy is then $\frac{1}{2}mv^2 = mgs$. But its

potential energy has been reduced to $mg(h - s)$, since $h - s$ is now its height above the point from which it was lifted. The sum of mgs and $mg(h - s)$ is mgh, the original potential energy. Whatever, then, the weight gains in kinetic energy as it falls, it loses in energy of the potential form. When the weight reaches the ground, the velocity acquired is given by the relation $v^2 = 2gh$, and therefore $\frac{1}{2}mv^2 = mgh$, or the potential energy possessed by the weight at the elevation h. During the fall the potential energy is continuously converted into the kinetic form, but in such a way that the sum of the two is a constant.

The heat and light produced by combustion are forms of kinetic energy, derived by transformation from the potential energy of chemical separation and chemical affinity.

When a storage battery is charged, the energy is stored as the potential energy of chemical separation. When it is discharged, the kinetic energy of the current flowing is derived by transformation from the potential form.

(f) *Thermodynamics.*—The science of thermodynamics treats of the laws governing the quantitative transformations of heat energy into other forms of energy. In other words, it is the science of heat considered as a form of energy, and includes the study of the relation of heat energy to all other forms of energy, whether mechanical, electrical or chemical. It is essential that the student of physical chemistry have a clear understanding of the fundamental principles of thermodynamics, since thermodynamics constitutes the very foundation of the entire science of physical chemistry.

(g) *The First Law of Thermodynamics.*—As has already been pointed out, we have no means of knowing what is the total energy in any one system, but when any change takes place in such a system there is usually a corresponding change in its total energy, either a decrease due to its giving up some of its energy to the surroundings, or an increase due to its receiving energy from the surroundings. These changes in the total energy of a system, we are able to study and to measure; and universal experience has demonstrated the fact that they obey the following law, known as the law of conservation of energy or the "first law of thermodynamics": "When a quantity of energy disappears at any place, a precisely equivalent quantity appears at some other place or places; and when a quantity of energy disappears in any form, a precisely equivalent quantity simultaneously appears in some other form or forms." In other words, energy can neither be

created nor destroyed. This fundamental principle was first clearly formulated by Mayer in 1842, following the publication of the results of the important investigations of Helmholtz, Joule, Hirn, and others, on the transformation of heat into work and of work into heat. These experiments established the fact that when work is transformed completely into heat, or heat into work, the quantity of work is equivalent to the quantity of heat.

(*h*) *Free Energy and Bound Energy.*—In 1882 Helmholtz introduced the terms "free" and "bound," to denote, respectively, the part of the kinetic energy of a system that is free or available for conversion into work, and the part that is not free or available for this purpose. It is evident that only that part of the energy contained in a system which can do work is of value. For example, in the case of a steam engine, the part of the energy given out by the fuel that is available for work is given by the ratio of the difference of temperature between the boiler and condenser to the absolute temperature of the latter; this means, of course, that only a part of the heat energy given out by the burning coal can be utilized, even in the most perfect engine.

A further important fact, also arising from experience, is that all other forms of energy may be converted without loss into heat, but the converse is not true. Heat cannot be converted completely to other forms except under well-defined conditions. This difference is due to the fact that in the case of heat energy the motion of the moving particles is unordered or random motion, whereas, in the case of all other forms of energy involved in any process of energy transformation, the movement of the parts whose motion gives rise to the energy in question is "ordered" motion, that is, it has only one, or at least a few definite directions, instead of having the random character of heat (molecular motion).

(*i*) *The Second Law of Thermodynamics.*—The first law of thermodynamics states that it is impossible to create energy; it is evident, therefore, that perpetual motion is impossible. We might, of course, conceive of a machine that could convert the heat of surrounding objects into mechanical work at the same temperature as itself. This would evidently be a perpetual motion; but since it differs from the first kind, it has been called perpetual motion of the second kind. However, experience teaches us that perpetual motion of the second kind is also impossible. In brief, then, heat cannot flow from a colder to a warmer body, nor even from a body of a given temperature to

another body of the same temperature. This is one of the many human experiences which led up to the formulation of the second law of thermodynamics.

The second law of thermodynamics deals with the "quantitative relations which restrict the convertibility of energy." It is somewhat abstruse and consequently has been stated in several different forms. The following three ways of stating it bring out different important points:

Clausius stated it as follows: "It is impossible for a self-acting machine, unaided by any external agency, to convey heat from one body to another at a higher temperature."

Nernst expressed it in this form: "While external work and the kinetic energy of moving bodies can be transformed into one another completely and in many ways, and can also be converted into heat, as by applying brakes to a railway train in motion, the reverse change of heat into work is only possible under certain conditions."

Lord Kelvin states the principle of the law in these words: "It is impossible, by means of inanimate material agency, to derive mechanical effect from any portion of matter by cooling it below the temperature of the coldest of the surrounding objects."

These statements apply only to the performance of devices working in cycles. Without this limitation, it is evident that the heat of a body, that of condensed gas, for example, may be converted into work by cooling the body below the temperature of surrounding objects; but before the operation can be repeated, the working substance must be restored to its initial condition, and this can only be done by applying energy from without.

Another deduction from the second law of energetics is that, "the cause of all action, or change, is the tendency of energy to obtain the same uniform degree of intensity as its environment." Moreover, the degradation of energy follows the line of least resistance. This is known as the "law of least action," and it is of great physiological importance, as will be pointed out later on.

(*j*) *Chemical Energy.*—It is a matter of everyday experience that there is some force which causes substances to combine and which holds them together, more or less firmly, in compounds. This force we call chemical energy, and all chemical reactions are manifestations of it. Moreover, these chemical changes are usually associated with the development or the absorption of heat. Substances that enter

readily into chemical combination usually give out a large amount of heat (for example, the combination of hydrogen and chlorine to form hydrochloric acid); but when combination is less vigorous, the heat given out is usually much less, and, in fact, heat may be absorbed in a chemical change.

We owe much of our knowledge of chemical energy to the French chemist, Berthelot; hence, a few of the fundamental principles or generalizations at which he arrived will be quoted here:

I. "The thermal change in a chemical reaction, if no external work is done depends only on the condition of the system at the beginning and end of the reaction, and not on the intermediate conditions."

II. "The heat evolved in a chemical process is a measure of the corresponding chemical and physical work."

III. "Every chemical transformation which takes place without the addition of energy from without tends to form that substance or system of substances, the production of which is accompanied by the evolution of the maximum amount of heat."

This third principle, which has come to be known as the law of maximum work, has also been stated by Berthelot as follows:

"Every chemical change which is accomplished without a preliminary action, or the addition of external energy, necessarily occurs if it is accompanied by disengagement of heat." This principle, especially when expressed in the latter form, is known as the "principle of the necessity of reactions." While there are many exceptions to this principle, it is, nevertheless, an important generalization.

(k) *Mass Unchanged in Chemical Reactions.*—It is very important to remember that one of the fundamental facts of chemical science is the conservation of mass. When a chemical reaction takes place, the substances change most of their properties: a solid may become a liquid or a gas; a liquid may become a gas or a solid; a poisonous gas like chlorine, may combine with a metal like sodium and form a highly nutritious substance (sodium chloride), and so on throughout the entire list of properties, all of which change except mass.

In 1895 Landolt conducted an elaborate investigation to determine whether the weight of the products of a reaction is exactly equal to the weight of the constituents before the reaction, weight being the accurate means of measuring mass. The result showed that in some cases slight differences existed between the products of reaction and the weights of the constituents before the reaction, but these differences

were in almost every case small enough to be attributed to error of technique.

The conservation of mass, then, is the one unchangeable property that matter exhibits when it is undergoing chemical transformations. It is on this fact that the science of quantitative chemical analysis is based.

(*l*) *Energy Unchanged in Chemical Reactions.*—It is an established experimental fact that the total energy is the same after a chemical reaction as before. However, it is rarely possible to convert *all* of the energy into another form. For instance, if before the reaction the total energy was in the form of chemical or intrinsic energy, after the reaction a part still remains in the form of intrinsic energy, and a part is transformed into heat; and, if there is a change in volume, as almost always occurs, a part is spent in doing external work. In brief, then, the law of conservation of energy holds absolutely in thermochemical changes, though it is seldom that *all* of the energy is transformed into heat.

An important corollary of this principle was announced by Hess in 1840. He advanced the following law, which he deduced from experimental data:

"The amount of heat generated by a chemical reaction is the same whether it takes place all at once or in steps." In other words, "all transformations from the original state to the same final state liberate the same amount of heat, irrespective of the process by which the state is reached." This implies, of course, that the final product or products will be the same in each case. Thus, when 12 grams of carbon combine with 32 grams of oxygen with formation of carbon dioxide, 94,300 calories are always liberated, quite independently of the rate of combustion, or the nature of the intermediate products.

This law is of the greatest importance for the indirect determination of the heat changes involved in certain reactions which cannot be carried out directly, such as the burning of foodstuffs in our bodies. Or, to give a concrete example, we cannot determine directly the heat given out when carbon combines with oxygen to form carbon monoxide. The heat given out when 12 grams of carbon are burned to carbon dioxide, as we have just seen, is 94,300 calories, which, by Hess' law, is equal to that produced when the same amount of carbon is burned to carbon monoxide and the latter then burned to carbon dioxide. The latter change gives out 68,100 calories; the reaction $C + O = CO$.

must therefore be equivalent to the liberation of 94,300 − 68,100 = 26,200 calories.

What are the factors of chemical energy? We have already learned the general principle of energetics, that every form of energy can be shown to be the product of two factors,—an *intensity* factor and a *capacity* factor, heat energy being the product of temperature and quantity of heat; electrical energy, the product of the quantity and the potential, etc. Applying the same principle to chemical energy, we must conclude that, if chemical energy can be resolved into two factors, one factor must be analogous to the capacity of thermal energy, and the other to the intensity factor. Willard Gibbs calls the intensity factor of chemical energy the *chemical potential*, in order to avoid the vagueness of the old term *chemical affinity*, which without doubt is the correct designation for "chemical intensity." Now, the quantity of a substance that takes part in any chemical change is proportional to the "equivalent weight" of the substance; and, assuming that the chemical equivalent is the capacity factor of chemical energy, we may postulate as follows:

chemical energy = equivalent weight × chemical potential.

It is fair to assume that the molecules of every substance possess a specific amount of chemical energy, which has a definite intensity under certain specified conditions. One substance can only react with another when the intensity of the energy associated with the original mixture is greater than that of the final system. If the intensity of the energy associated with the original mixture is the same as that associated with the products of the reaction, no reaction will take place; if the intensity factors are not equal, the energy will not usually be at rest. Chemical reaction is, then, in a sense, analogous to the flow of electricity from a higher to a lower potential, or to the flow of heat from a higher to a lower temperature.

Ostwald has drawn attention to the fact that, if the chemical process is performed in a battery cell, the work derived from the process will be transformed into an equivalent amount of electrical energy without passing through the form of heat. And, since by Faraday's law the capacity factor—quantity of electricity—is proportional to the quantity of matter decomposed, the capacity factor of the electrical energy will be proportional to the capacity factor of the chemical energy. Hence the respective intensity factors of chemical and elec-

trical energies will also be proportional. But electromotive force is proportional to the intensity factor of electrical energy, consequently the "electromotive force is proportional to chemical affinity." Or, in the words of Faraday, "the forces called electricity and chemical affinity are one and the same." Chemical action, therefore, takes place when the potential of the reacting substances is greater than that of the reacting products, and we can consider the "affinity" between reacting substances roughly in terms of difference of potential.

3. The Storage of Potential Energy in Living Matter: The Photosynthesis of Carbohydrates.—It has already been pointed out that living matter is a machine in which chemical and other forms of potential energy are transformed into various forms of kinetic energy, or into the chemical energy of new compounds.

The kinetic energy of living matter may appear as heat, as mass movements, as light, or as electrical energy. Experience shows that chemical transformations of living things are for the most part exothermic, since living things constantly produce heat, and their temperature is usually a little above that of their environment. In some cases, however, the energy appears as light rather than heat. This is the case, for example, in the luminous organs of the firefly.

Electrical changes occur in all cells when combustion takes place in them; in fact, at times nearly the whole of the energy may appear in this form, as, for example, in the electrical organ of the torpedo.

In brief, then, the phenomena that are peculiarly characteristic of vital changes are those associated with the transformation of energy; and, since energy can neither be created nor destroyed in living matter any more than it can in the inorganic world, it is evident that all life processes demand for their continuance and maintenance a continuous supply of energy from the *outside*.

Careful experiments prove that all the energy of living matter has come in the long run from the sun in the form of light energy. This is very important: *light*, not heat, is the synthetic agency that brings about the synthesis of living matter. Sunlight, with its potential energy acting upon chlorophyll-containing plants, synthesizes organic compounds (carbohydrates, etc.) from water, carbon dioxide, etc. In the words of Moore: "The higher plant takes up the radiant solar energy, and, from inorganic materials, the most important of which are carbon dioxide, water, and nitrates, it builds up organic bodies, containing chemical energy, for its own purposes in growth, reproduc-

tion, increase. These synthesized products yield directly or indirectly the energy for all the higher animals, in whose bodies, by up-and-down processes of oxidation and reduction, but on the whole with a downward trend of oxidation, the synthesized products are converted again into simple inorganic bodies, carbon dioxide, water, and very simple nitrogenous compounds. Thus the great cycle of Nature is completed. Those plants which do not contain chlorophyll, and hence cannot synthesize directly from inorganic materials, are compelled, like animals, to live on other plants or decaying vegetable matter, or upon the bodies of living or dead animals."

It is evident that all living matter may be divided into two classes: (a) that which contains a green pigment called chlorophyll, and which in consequence is capable of acting as the transformer of light energy into chemical energy, etc.: and (b) that which is dependent upon other organisms, either living or dead.

The Photosynthesis of Carbohydrates.—The simplest carbohydrates known to occur abundantly in plant tissues are the hexoses, having the formula $C_6H_{12}O_6$, which is just six times that of formaldehyde, CH_2O. Also, it is known that formaldehyde easily, and even spontaneously, polymerizes into more complex forms having the general formula $(CH_2O)_n$; e.g., trioxymethylene, $C_3H_6O_3$. Further, both formaldehyde and trioxymethylene can easily be condensed into hexoses by the use of suitable catalytic agents. Hence, it is commonly believed that formaldehyde is the first synthetic product resulting from photosynthesis,[1] that this is immediately condensed into hexose sugars, and that these in turn are united into the more complex carbohydrates.

The leaves and other chlorophyll-containing tissues of green plants absorb carbon dioxide from the atmosphere and simultaneously liberate an equal volume of oxygen. Within certain limits, the rate at which this reaction takes place is proportionate to the intensity of the illumination of the leaf, and to the percentage of carbon dioxide in the atmosphere. Not all parts of the spectrum are equally efficient in promoting this process, the red rays between B and C causing the most rapid assimilation, while the activity of the rays between D and E on the Fraunhofer scale is a minimum, and there is a second maximum in the violet beyond F.

---

[1] Photosynthesis is the term applied to the process whereby chlorophyll-containing plants, in the presence of sunlight, synthesize organic compounds from water and carbon dioxide, etc.

When the last-mentioned form of radiant energy (known as ultra-violet light, i.e., light of very short wave-length, beyond the spectrum visible to the human eye) is present is large amounts, chlorophyll is not needed as a transformer in the production of formaldehyde from carbon dioxide and water, as is evidenced by the following experiment: If a mixture of CO_2 and water is illuminated by the light of a mercury arc in a tube of pure silica, this light being particularly rich in ultra-violet rays and silica not absorbing them as glass does, the mixture is very soon found to contain hydrogen, oxygen, carbon monoxide, carbon dioxide, formaldehyde, and hydrogen peroxide. The reaction goes to the point of equilibrium; it is never complete. If one starts with a carbohydrate solution, or with formaldehyde and oxygen, the reaction is reversed to the same mixture as if one started with carbon dioxide and water. The reaction may be written thus:

$$\text{light} + 2CO_2 + 2H_2O \rightleftarrows O_2 + CO + CH_2O + H_2O_2.$$
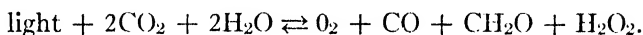
The energy of ultra-violet light has been absorbed and appears as the potential energy of the oxygen-formaldehyde system.

In ordinary life processes, however, but little energy is obtained from this source, because the amount of ultra-violet light which reaches the earth's surface after passing through the atmosphere is small. The advantage of chlorophyll is that it makes possible the utilization of the longer radiant-energy vibrations, which reach the earth in greater abundance, or with more energy in them. Chlorophyll absorbs the red and orange vibrations and hence has a green color. The energy of the absorbed red light is then used in the synthesis of formaldehyde, just as ultra-violet light is used in the absence of chlorophyll.

The mechanism for converting solar radiation into bound energy is imperfectly known at present; however, certain facts have been ascertained:

(1) Elements taken from the air and soil are undoubtedly built up into organic compounds almost entirely in the presence of sunlight. It has been shown, for example, that during the day the starch content of the leaves of certain plants may rise to 6.44 per cent of the dry leaf weight, whereas at night the starch content may drop as low as 0.38 per cent. Moreover, it has been shown that starch formation is greatest where there is greatest absorption of light. This fact was demonstrated by Timiriaff in the following manner: He threw a spectrum upon a starch-free leaf and then determined the relative amounts

of starch produced under different spectral regions. His results showed that the greatest activity lies where most light is absorbed by a chlorophyll solution, viz., in the red between the B and C lines. Other carbohydrates are also found in variable quantities.

(2) It has been found that carbon dioxide is absorbed beyond the needs of respiration, and that oxygen is evolved. The following experiment of Englemann shows that the maximum evolution of oxygen takes place where there is greatest absorption of light, which, as stated above, is where the maximum production of starch takes place. Englemann placed a filament of Cladophora in water, to which he added some motile bacteria having an avidity for oxygen. On the thread of algæ he projected a minute solar spectrum and kept it under the microscope. It was observed that the bacteria gathered just at those places (red and violet) where light was absorbed.

Kniep and Minder have estimated the carbon assimilation, and they find it to be directly proportional to the amount of energy absorbed as light. This fact seems to suggest that the actual pigment itself is merely an optical sensitizer, since there is no relation between its particular absorption bands and the photo-chemical change.

It is fair to conclude, then, that the maximal effect of the chlorophyll system is in relation to that part of the spectrum which is absorbed most.

Von Baeyer, in 1870, advanced the hypothesis that formaldehyde, formed by the reduction of carbon dioxide, is the first product of assimilation in plants, and that this substance subsequently undergoes polymerization to carbohydrate. Proceeding upon this assumption, many investigators have studied the question as to whether formaldehyde is actually produced in green leaves. Some reported in the affirmative, whereas others have criticized these results, claiming that the experiments were inconclusive.

However, more recent investigators have been more successful. Usher and Priestley, during the years 1906 to 1911, made an elaborate study of the subject. Briefly, their results showed that the photolysis of carbon dioxide and water results in the formation of formaldehyde and hydrogen peroxide, as represented by the following equation:

$$CO_2 + 3H_2O = CH_2O + 2H_2O_2.$$

The formaldehyde is then condensed by the protoplasm into sugars, while the hydrogen peroxide is decomposed by an enzyme in the plant

cell, into water and oxygen. If the formaldehyde is not used up rapidly enough by the protoplasm, it kills the enzyme, and the undecomposed hydrogen peroxide destroys the chlorophyll, which stops the whole photosynthetic process. These investigators were also able to cause the photolysis of carbon dioxide and water into formaldehyde outside of a green plant, by the use of a suitable catalyzing agent which continually destroys the hydrogen peroxide as fast as it is formed. They showed the actual bleaching effect of an excess of hydrogen peroxide in plant tissues which had been treated in such a way as to prevent the enzyme from decomposing it. Moreover, they were able to demonstrate the condensation of formaldehyde into starch by the action of protoplasm which contained no chlorophyll.

In the mean time, definite proof of the formation of formaldehyde was given by Fenton (1907) who showed that it is formed when carbon dioxide is reduced by means of metallic magnesium.

This observation of Fenton assumes greater importance when it is considered in conjunction with Willstätter's recent discovery that chlorophyll contains magnesium as an integral part of the molecule. He regards the magnesium as playing just as important a rôle in the process of assimilation in plants as does the iron content of hæmoglobin in its function as oxygen carrier.

More recently, Ewart has shown that formaldehyde can combine chemically with chlorophyll; from this fact, Schryver deduces the theory that if for any reason the condensation of formaldehyde into carbohydrates by the cell protoplasm does not take place as rapidly as the formaldehyde is produced by photosynthesis, the excess of the latter enters into combination with the chlorophyll; and that, if condensation into carbohydrate uses up all the free formaldehyde that is present in the active protoplasm, the formaldehyde-chlorophyll complex is broken down, setting free an additional supply for carbohydrate manufacture. According to this conception, there are in the chloroplasts (the chlorophyll-bearing protoplasm), not only the agencies for the production of formaldehyde from carbon dioxide and water and for the condensation of this into carbohydrates, but also a chemical mechanism by means of which the amount of free formaldehyde in the reacting mass may be regulated so that at no time will it reach the concentration which would be injurious to the cell protoplasm, or fall below the proper amount for carbohydrate formation. This explanation affords a satisfactory solution of the difficulty encountered by the

earlier investigators of photosynthesis, namely, the fact that free formaldehyde is powerfully toxic to cell protoplasm. It also explains why it was so difficult to prove the presence of formaldehyde in plant cells. Formaldehyde does not accumulate in the cell, since it is withdrawn to form sugar as fast as it is formed.

It is worthy of note that chlorophyll, which is without doubt the principal agency acting as a transformer of light energy into the potential energy of living plants, is one of the most highly organized and complex substances in living matter, and in consequence cannot be thought of as being the first stage in the evolution of organic from inorganic matter. The protoplasm of the living cell is likewise too complex to be considered in this connection; consequently, many investigators have sought among the inorganic compounds for some substance or substances having synthesizing powers. Experiments with crystalloidal inorganic matter have met with doubtful success; but Moore and Webster and others have definitely proved that inorganic colloidal uranium, iron, and aluminum hydroxides possess the power of transforming the energy of sunlight into the chemical energy of organic compounds. Moreover, it has been found by Moore and his colleagues (a) "that inorganic iron, in crystalloidal or colloidal form, is present in the colorless part of the chloroplasts of the green plant cell in many plants; (b) in the absence of iron the green coloring matter cannot develop in the leaf, although the green coloring matter itself contains no iron; (c) in the presence of sunlight the iron-containing substance of the chloroplast develops the coloring matter, so that this itself is a product of photosynthesis induced by the iron-containing compound." In other words, there is evidence to show that inorganic colloidal systems were the precursors of our present-day chlorophyll system, since these substances seem to be possessed of the power of fixing carbon and nitrogen and building these up into organic compounds with the uptake of energy from the sunlight. The inorganic colloidal systems, however, seem to be able to utilize only light of short wave-length, whereas the chlorophyll, which acts as the transformer in the highly organized cell, utilizes the more abundant longer wave-lengths, and in addition, by means of its green color screen, protects the delicate plant cells from the short wave-lengths, which are harmful to them.

Fats are also stored up in plants. Just how they are synthesized is not known. Whether the plants can form them directly or whether

they are only synthesized from carbohydrates has not been determined. It is known, however, that they can be formed from carbohydrates; moreover, it is known that this action is exothermic.

Proteins also contain bound energy, which comes indirectly from the sunlight. For instance, certain bacteria have the power of "fixing" atmospheric nitrogen in a form available for the use of plants. The nitrogen so fixed carries with it a considerable quantity of energy.

To sum up, photosynthesis only occurs in the chlorophyll-containing parts of the plant, and in these only when they are exposed to light. Incidentally, man obtains the energy necessary for his maintenance and for the performance of physical work from the disruption of carbohydrates, fats, and proteins, which were synthesized in the first instance by green plants through the agency of the sun's energy. In other words, the plant acts as a transformer of kinetic energy into potential energy by the formation of carbohydrates, fats, proteins, and other essential foodstuffs.

4. Energy Transformations in Animals.—(*a*) *Energy Relations of Plants and Animals.*—We have learned from the foregoing paragraphs that, broadly speaking, the constructive chemical processes in nature are the work of the higher plants. Animals, as well as those forms of plant life that lack chlorophyll, carry on for the most part reactions of a destructive character. Early investigators, however, were wont to exaggerate the contrast between the chemical reactions of plant cells and those of animal cells. They were disposed to discredit the ability of animal cells to synthesize new compounds. This was an erroneous conception, for while it is true that animal cells do not make any direct use of the energy of sunlight in the manufacture of compounds, nevertheless, they do carry on synthetic reactions, though to a lesser extent than do plants.

It is, of course, a well-known fact that plant life and animal life are antithetical and at the same time complementary to each other. Animals breathe in oxygen and exhale carbon dioxide; while plants use the carbon dioxide of the air (which is largely that exhaled by animals) as a part of the raw material used in photosynthesis, and exhale oxygen. Plants absorb simple gases, such as carbon dioxide and ammonia from the air, mineral substances from the soil, and water from both the air and soil, and build these up into complex carbohydrates, fats, proteins, etc.; while animals use these complex com-

pounds of plant origin as food, transforming parts of them into various other forms of structural material, but in the end these are again broken down into simple gases and mineral substances, which are eliminated from the body.

However, many of the compounds produced in plant and animal cells are identical in structure, undergo similar chemical changes, and are acted upon by similar catalysts. Plant cells exhibit respiratory activities—use oxygen and give off carbon dioxide in exactly the same way that animal organisms do. We have learned that the constructive photosynthesis of plants is effected through the aid of chlorophyll. This pigment is remarkably similar to the blood pigment, hæmatin, which controls the vital activities in the animal organism, differing from the latter only in the mineral element which links the characteristic structural units together in the molecule. There are innumerable other points of similarity in the chemistry of the vital processes of plants and animals.

(b) *Calorimetry.*—Turning now to the definite problem of animal energy transformations, let us consider the methods of measuring the potential energy of the various foodstuffs, and compare the energy value so found with the actual amount of energy liberated in the organism. The fundamental facts will not require detailed consideration. Living animals in their various energy changes obey the laws of energetics. The energy of the income is potential in the complex molecules of the food. It is released in the oxidative decomposition processes of animal life, and made kinetic. The actual fuels burned by the animal machine are carbohydrates, fats, proteins, alcohols, etc.; and it is evident that we are concerned with the *heat of combustion* of these substances in energy measurements of metabolism.

The unit of heat adopted by biologists is the large Calorie—that is, the quantity of heat required to raise the temperature of 1 kilogram, (1 liter) of water 1° C., say, from 20° to 21°. (The small calorie is $\frac{1}{1000}$ of the large.) To avoid confusion, the large Calorie is invariably distinguished from the small by the capital C.

It would be entirely out of place in this text to enter upon a discussion of the general problems of metabolism; we shall confine ourselves to the following problem: Does the heat energy given out by the animal body equal that which would be given out by the complete combustion of the various foodstuffs outside the body? Or, to be more exact: Does the heat of combustion of a foodstuff, as determined in a

"bomb calorimeter," equal that produced in the animal body, *plus* that given out by the complete combustion of the excreta?

The general problem in *thermochemical measurements* is to determine the amount of heat that is liberated in chemical reactions, such as oxidation. In order to accomplish this by direct measurement, the heat that is set free by the reaction is allowed to warm a known quantity of some liquid (usually water) whose specific heat is known. The rise in temperature is then measured by means of a very accurate thermometer, preferably a Beckmann thermometer (see p. 140).

If we had to deal only with simple chemical compounds in metabolic processes, it would not be necessary, as a rule, to use mechanical instruments to get the calorific value of the substance in question, since the energy of a pure chemical compound may be calculated from its chemical formula; but the common foodstuffs and the excreta, which are the types of substances in which we are interested, are very complex mixtures, hence this procedure is impracticable.

The heat of combustion of any substance, such as the foodstuffs, is usually determined by burning a known weight of the substance in question in a *bomb calorimeter* in an atmosphere of oxygen, and then measuring the rise in temperature of the known amount of water in which the bomb is immersed. The rise in temperature multiplied by the weight of the water gives the amount of heat expressed in Calories.

Fig. 9.—The Bomb of a Bomb Calorimeter.
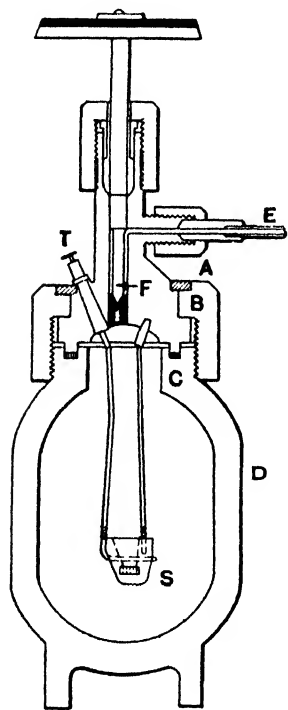
The Bomb.—There are many different forms of bombs on the market, but they are all fashioned, more or less, after the original Berthelot type. A convenient form is shown in Fig. 9. It is known as the Mahler-Cook bomb, and is a modification of the Berthelot-Mahler bomb.

The bomb consists essentially of an enamel-lined steel vessel, capable of withstanding high pressures.

The lower part, D, is closed by a lid, A, which is screwed on, an air-tight connection being obtained by means of a lead washer, C. Two stout platinum wires pass through the cover, one, T, being insulated by means of a quartz plug. These wires are connected by two terminals. One of the wires is bent round in the form of a loop, so as to support a crucible, which may be of platinum, unglazed porcelain, or silica. The substance is placed in the crucible, and the ignition is effected by means of a coil of iron wire, which joins the two platinum wires, and which is caused to burn by means of an electric current. Oxygen is introduced through a valve in the center of the lid, the opening and closing of the valve being controlled by screw F. The oxygen, which is supplied from a cylinder fitted with a pressure gauge, is connected at E.

FIG. 10.—A Bomb Calorimeter

The Calorimeter.—The calorimeter consists of a large nickel-plated vessel (see Fig. 10), containing the water in which the bomb is immersed. This is surrounded by an outer water-jacketed vessel. Both calorimeter and outer vessel are fitted with appropriate stirring arrangements. The stirrer in the calorimeter should be worked with a small motor. The stirrer should be so arranged that it will almost touch the bottom of the calorimeter on the downward stroke, and it should remain completely immersed in the water on the upward stroke.

The change of temperature should be obtained by means of a Beckmann thermometer. The outer jacket is closed with a non-conducting lid, fitted with the necessary holes for the thermometer and stirrer.

Any given oxidation will always produce the same amount of heat. Thus, if we oxidize a gram of carbon, a known amount of heat is produced, whether the element is free or in a chemical compound. The following figures show the approximate number of heat units produced by the combustion of one gram of the following substances:
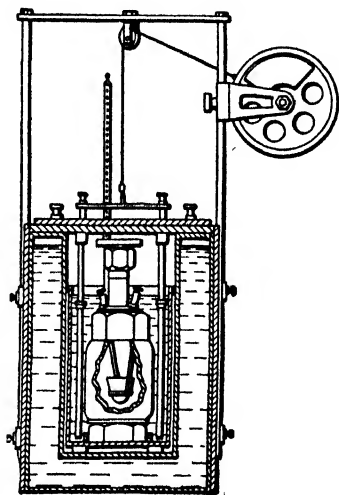
	Cals.		Cals.
Hydrogen	34.66	Tissue-fat	9.48
Carbon	8.10	Butter-fat	9.23
Urea	2.53	Olive oil	9.33
Alcohol	7.00		——
		Average for fats	9.35

Proteins:	Cals.	Carbohydrates:	Cals.
Casein	5.86	Glucose	3.74
Egg-albumen	5.74	Sucrose	3.96
Serum-albumen	5.92	Lactose	3.95
		Starch	4.18
	——		——
Average	5.84	Average	3.96

It is customary in metabolic experiments to use, for carbohydrates and fats in a mixed diet, the figures that were obtained by Rubner, namely:

One gram of carbohydrate = 4.1 Cal.
One gram of fat   = 9.3 Cal.,

the high value for carbohydrates being due to the predominance of starch in the usual diet.

The actual heat production observed when carbohydrates, fats, or alcohol (the Cal.value of alcohol is 7.0) are burned in a bomb calorimeter, or inside the body by the metabolic processes, is identical in both cases, since the end products are the same simple products, namely, carbon dioxide and water. The *proteins,* however, stand somewhat apart in their behavior, owing to the fact that they are not completely burned in the animal body, the nitrogen being excreted as urea, creatinine, etc., substances which have a certain capacity for oxygen. Urea, for example, when burned in a bomb calorimeter, has a fuel value of about 2.53 Cal. per gram. The residual fuel value in urea is an indication that a certain amount of the energy stored up in proteins is constantly lost to the animal economy. Bacteria may profit by it, but it seems to be unavailable to the higher forms of life. Furthermore, proteins as they actually occur in the diet are not completely digested or assimilated, and the undigested or unassimilated material is eliminated in the feces.

To make accurate allowance for the energy lost with these incompletely oxidized compounds is a difficult matter. Rubner was the first investigator who was successful in determining the actual calorific value for protein in the animal body. His method was a follows: He

fed a dog on washed meat, and subtracted from the heat of combustion of the food (determined in a combustion calorimeter) the heat of combustion of the urine and feces, plus a small correction representing the heat of solution of the urea in the urine. From these data it was found that an average of about 4.1 Calories per gram were actually available to the animal from the protein in the diet. In brief, then, the actual calorific value of a protein to an animal is about the same as that of a carbohydrate.

It is obvious that we can compute the total available energy of a diet by multiplying the quantity, expressed in grams, of each foodstuff by its Calorie value.

The Respiration Calorimeter.—A direct measurement of the energy that is actually liberated in the animal body can also be obtained by the use of a special type of calorimeter known as a "respiration calorimeter." This, of necessity, must be of somewhat different construction from the bomb calorimeter, for it must provide for long-continued observations and for an uninterrupted supply of oxygen to the animal. It must also be provided with means for the measurement of the amounts of carbon dioxide and water discharged and of oxygen absorbed by the animal during the observation. Many respiration calorimeters have been made for all sorts and sizes of animals; but by far the best is the Atwater-Rosa-Benedict instrument, which is of sufficient capacity to hold a man, a table and chair, a bed, a telephone, an appliance for the performance of mechanical work (a bicycle attached to an ergometer), etc., and yet so technically perfect that when a measured amount of heat was generated within the calorimeter by an electric current the quantity of heat liberated could be measured to within 0.01 per cent. Moreover, the actual heat given out when a definite amount of ether or alcohol is burned in the respiration calorimeter exactly corresponds to the amount obtained when these substances are burned in a bomb calorimeter.

All the energy liberated in the body does not, however, take the form of heat. There is a variable amount which appears as mechanical work, which must be added to the actual calories given out in order to measure all of the energy which the animal expends. We can express this work in terms of heat units, however, just as we can express heat in terms of electrical units, or electrical units in terms of mechanical work. Since no energy is ever lost and all forms of energy are equivalent to one another, the heat value consumed in performing mechanical

work can be directly calculated from the known mechanical equivalent of heat. (A Calorie corresponds to 425 kilograms-meters of work.)

As illustrating the extreme accuracy of the respiration calorimeter, take the following actual experiment: A man doing work on a bicycle gave out as actual heat 4833 Cal., and did work equivalent to 602 Cal.,

Window Window
Desk Table Telephone
Balance
Window Galvanometer Water Meter
Ingoing Air Door
Outcoming Air
Observer's Table H₂O Absorbers
Rotary Blower Balance
Circuit Breaker Respiration Chamber Water Pressure Regulator
Shaft CO₂ Absorbers Alcove
Pan Pan Rack
Table Table
Absorption System
Window Alcove
Sink
Water Pump Tool Closet
Electric Motors Barometer
Balances Drying Ovens
Window Door

FIG. 11.—The General Appointments of a Respiration-calorimeter Laboratory.

giving a total of 5435 Cal. The energy of the food consumed by the subject during the period of the experiment was capable of yielding 5459 Cal., which may be considered as exactly balancing the actual output.

It would be out of place in this text to give a detailed description of the respiration calorimeter. A fairly accurate idea of the general

construction and appointments of the Atwater-Rosa-Benedict apparatus may be obtained from the two accompanying figures.   Figure 11 gives one an idea of the general appointments of the respiration-calorimeter laboratory, and Fig. 12 explains the principle by which the output is measured.

On referring to Fig. 12 one can see that the calorimeter is connected with a closed system of tubes provided with an air-tight rotatory blower or pump to maintain a constant current of air, as indicated by
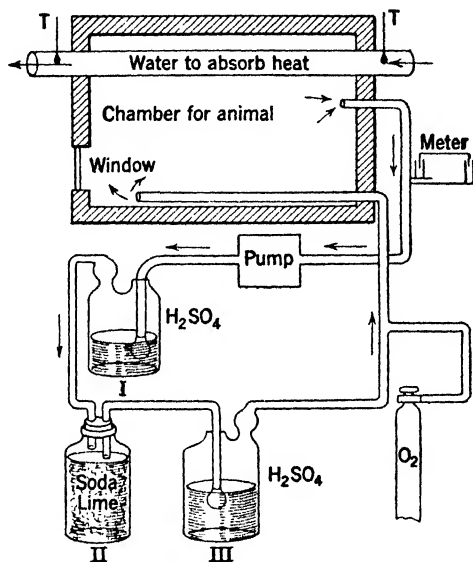
FIG. 12.—Diagram of the Atwater-Rosa-Benedict Respiration Calorimeter. As the animal uses up the O_2, the total volume of air shrinks. This is indicated by the meter, and a corresponding amount of O_2 is delivered from the weighed O_2 cylinder. The increase in weight of bottles II and III gives the CO_2; that of I, the water vapor.

the arrows. Following the air stream as it leaves the chamber, one notes a side tube connecting with a meter to indicate changes in volume of the air in the system. Beyond this and the pump is a specially constructed bottle containing concentrated H_2SO_4, then one containing soda-lime, and finally another H_2SO_4 bottle. The first H_2SO_4 bottle absorbs all the water vapor contained in the air coming from the chamber; the soda-lime bottle absorbs the CO_2; and the second H_2SO_4 bottle absorbs the water that is produced in the chemical reaction involved in the absorption of the CO_2 by the soda-lime ($2NaOH + CO_2$

= H_2O Na_2CO_3). By weighing these absorption bottles before and after an animal has been in the chamber for some time, the weight of H_2O and CO_2 given out can be determined. Another side tube leads to an oxygen cylinder, the valve of which is so regulated as to cause oxygen to be discharged into the system at such a rate as to compensate exactly for that used up by the animal, as indicated by the behavior of the meter. The amount of oxygen required may be determined by weighing the oxygen cylinder before and after the observation.

To complete the investigation, it is necessary that the urine and feces be collected and analyzed.

By the use of such delicate apparatus as the calorimeter just described, the various sources of error due to faulty technique have been eliminated, and the results that have been obtained by different investigators are so satisfactory as to convince even the most skeptical of observers that *the oxidation of the various foodstuffs in the living body produces the same amount of energy as would be produced if they were burned in a bomb calorimeter, provided the end products are the same.* Also, the law of conservation of energy is proved to be as true within the animal body as elsewhere in nature. As an energy transformer, the animal body behaves in all respects like a machine.

CHAPTER V

GENERAL NATURE OF SOLUTIONS

1. Importance of Solutions.—In the introductory chapter it was pointed out that the living cell is a complex solution containing both colloids and crystalloids and that the chemical reactions occurring in living cells are reactions in solution. In other words, we have evidence to show that protoplasm is a fluid mass consisting of a solution or suspension in water of compounds of decidedly varying complexity. Moreover, all the energies that are displayed in the living cell are undoubtedly derived from the energies of substances in solution. That is to say, in the life processes no material is available for energy purposes unless it is soluble, and until it is rendered soluble it cannot be absorbed or utilized. For example, until solution occurs, the foodstuffs cannot obtain entrance into a living organism; they cannot get through the walls of the intestine into the blood; and if they could get into the blood they could not get into the living cells. It is evident, then, that all the chemical changes that take place in the building up and breaking down of cells are changes in and between substances in solution. In the chapter on Energy Transformations in Living Matter the author spoke of the energy of the body as being derived from the combustion of foodstuffs. This statement is misleading if it conveys to the mind too strongly the phenomenon of burning fuel. It is true that the end products resulting from the oxidation (burning) of such substances as carbohydrates and fats are the same whether the process takes place inside the body or outside, i.e., CO_2 and H_2O, and the energy liberated is the same; but inside the body the process occurs by continual gradations, the reacting substances in every case being dissolved in water. In fact, as we shall learn later on, the process of oxidation in the body is largely a process of substitution of $-OH$ ions for H ions, and the reverse process is what is known as reduction under similar conditions.

In short, the available energy of a living cell at any given time may be regarded as made up of two factors, viz.:

(1) The chemical energy of the various substances that compose it, i.e., the energy which will be produced when the component parts are completely oxidized.

(2) The kinetic energy which the substances possess because they are in solution.

From the foregoing statements one can deduce that the mere solution of a substance may so alter the state of that substance as to set energy free.

What has been said about the importance of solution to living matter applies to other branches of science.

In chemical procedures our first problem is usually to get the material into solution, since practically all chemical reactions take place in solution. In physical chemistry, and even in pure physics, we likewise find solution playing an important part in the procedures. We may mention the phenomena of surface tension, viscosity, diffusion, polarization, electrolysis of homogeneous and heterogeneous mixtures, the combustion of heat in solutions, and the absorption of light by solutions, and the manufacture of primary and secondary batteries.

It is evident from the above remarks that our interest in solution will continue as long as the natural sciences are studied. To quote H. C. Jones of Johns Hopkins University: "In solution we are dealing not only with a fundamental condition of matter, but with *the fundamental condition.*"

2. Types of Solutions.—We have seen that matter occurs in nature in different states or conditions known as *states of aggregation.* Matter in any one of the three principal states—the gaseous state, the liquid state, and the crystalline state—tends to form more or less homogeneous mixtures with matter in the same state of aggregation as itself and with matter in either of the other states. When such a mixture is chemically and physically homogeneous, and no abrupt change in its properties results from a change in the proportions of the compounds of the mixture, it is a one-phase system and is termed a solution. The meaning of this definition will be seen from the following example:

If we mix very carefully some glucose, $C_6H_{12}O_6$, and sand, SiO_2, we can obtain a mixture of these two materials, but on careful examination of the product we can easily detect the presence of two crystalline substances, i.e., crystals of pure glucose and of pure sand. This mixture is, therefore, not a one-phase system, but is a heterogeneous mixture. On the other hand, when glucose is mixed with water, we

obtain, when equilibrium is established, a homogeneous, one-phase system composed of two different substances, uniformly mixed with each other. Such a system is called a true solution, to distinguish it from heterogeneous mixtures, such as colloidal solutions.

Mixtures of the types just mentioned may be classified as follows:

(1) solutions of gases in gases;
(2) solutions of liquids in gases;
(3) solutions of solids in gases;
(4) solutions of gases in liquids;
(5) solutions of liquids in liquids;
(6) solutions of solids in liquids;
(7) solutions of gases in solids;
(8) Solutions of liquids in solids;
(9) solutions of solids in solids.

While examples of all these different types of solutions are known, only a few of the more important of them will be considered in this chapter. Colloidal solutions (heterogeneous mixtures that approach true solutions in their properties) will be considered in a later chapter.

In the systematic treatment of solutions it is convenient to classify them into three groups: Group I, *gaseous solutions*, or solutions of gases, liquids, and solids, in gases. Group II, *liquid solutions*, or solutions of gases, liquids, and solids, in liquids. Group III, *solid solutions*, or solutions of gases, liquids, and solids, in solids.

It is evident that Group II is by far the most important of the three, because most chemical, many physical, and practically all biological processes are carried out in such solutions.

3. Solutions of Gases in Gases.—Of Group I the most important type is the solution of gases in gases. In solutions of this class, the components may be present in any proportion, since gases are completely miscible. If no chemical change takes place on mixing two gases, each gas behaves independently, and the properties of the mixture are the sum of the properties of the constituents. Thus, the pressure exerted by a mixture of gases is the sum of the pressures which each gas would exert if it alone occupied the total volume. This law was first discovered by Dalton, and is known as "Dalton's law of Partial Pressures." It holds with reasonable accuracy when the gases are not too concentrated, i.e., when the pressures are not great. At

great dilution this generalization is nearly true, but at high pressures many exceptions have been discovered.

Solutions of liquids in gases and of solids in gases will not be considered here. *Clouds* are a good example of the former type, and *smoke in air* of the latter.

4. Solubility of Gases in Liquids.—In contrast to the complete miscibility of gases in gases, liquids are only capable of dissolving gases to a limited extent, the amount dissolved depending upon the nature of the gas and the liquid solvent, as well as upon the temperature and pressure of the system.

When a liquid will not take any more of a gas at a given temperature, it is said to be saturated with the gas, and the resulting mixture is termed a *saturated solution.*

Some gases are only slightly soluble in a given solvent, whereas others are very soluble. For instance, one volume of water, at normal temperature and pressure, will dissolve 1200 volumes of ammonia, and but 0.021 volume of hydrogen.

Influence of pressure.—The greater the pressure of a gas, the greater the quantity of it taken up by the solvent. In 1803, Henry discovered an important relationship between the pressure and the solubility of a gas. Under a total pressure of 4 atmospheres, a saturated solution of gas holds in solution 4 times as much gas, by weight, as under a total pressure of 1 atmosphere. Now, according to Boyle's law, the concentration of a gas in an enclosed space is proportional to the pressure; i.e., by doubling the pressure the volume of gas is halved. Or, to use the quantities used above, 4 volumes of gas at atmospheric pressure occupy but 1 volume at a pressure of 4 atmospheres. From these data Henry formulated his law, which is: "Under equal conditions of temperature, water takes up in all cases the same volume of condensed gas as of gas under ordinary pressure."

With regard to the influence of the nature of the gas on the solubility, it may be said in general that the extremely soluble gases, like ammonia and hydrogen chloride, do not obey Henry's law in the matter of solubility, i.e., their solubility is not proportional to the pressure. This is undoubtedly due to their distinct basic and acidic properties, which induce them to form compounds with the solvent. Thus, NH_3 unites with $H \cdot OH$ to form NH_4OH, and HCl unites with $H \cdot OH$ to form a hydrate, $HCl \cdot H \cdot OH$.

For comparative purposes, the solvent power of a liquid for a gas is

usually expressed in terms of the "coefficient of solubility," i.e., the volume of the gas taken up by a unit volume of the liquid at a given temperature. The standard is the volume of a gas, reduced to 0° C. and 76 cm. pressure, which is taken up by a unit volume of a liquid at a definite temperature under a gas pressure of one atmosphere.

The influence of temperature on the solubility of gases in liquids is very definite, the solubility diminishing fairly rapidly with rise in temperature.

5. Solutions of Liquids in Liquids.—In dealing with the mutual solubility of liquids in liquids, two types are sharply distinguishable: (1) those that mix completely in all proportions, e.g., alcohol and water; (2) those that are partially miscible, e.g., ether and water. These two classes of solutions differ considerably in their properties. The second class, however, shows the greatest range of variability; some liquids show tendencies to complete miscibility, and other are very nearly insoluble. It must be remembered that there is no liquid which is absolutely insoluble in another liquid.

Completely Miscible Liquids: Class I.—The properties of liquids that mix in all proportions suggest the solution of two or more gases, with the distinction, however, that in this case the properties are *not* the sum of the properties of the constituents. The volume is never equal to the sum of the volumes of the constituents. Usually the volume decreases on mixing, but in some instances it increases. This difference is due to the mutual rearrangement of the molecules, and in some cases to the disruption of associated molecules.

There is always a change in temperature when liquids are mixed. In some cases heat is evolved, in others it is absorbed. As yet, no relation has been discovered between the changes in volume and the thermal changes which accompany the mixing of liquids.

There are some cases in which the properties of completely miscible liquids are nearly additive, and in such cases the properties can be approximately calculated before the liquids are mixed.

Partially Miscible Liquids: Class II.—Many liquids are known that are only partially soluble in each other. A mixture of ether and water is a good example of this class. One hundred grams of water at a temperature of 19° C. dissolve 6.86 grams of ether; and at the same temperature 100 grams of ether dissolve 1.21 grams of water.[1] On mixing equal quantities of water and ether in a separatory funnel, the

[1] Klobbie, Zeit. phys. Chemie, **24**, 619.

solution will at first appear to be homogeneous, but on standing two distinct layers will be observed. The top layer will be found to be a saturated solution of water in ether, and the lower layer a saturated solution of ether in water. The relative quantities present in each saturated solution will depend on the temperature of the mixture.

In most cases, the solubility of partially miscible liquids increases with rise in temperature, and in such cases there is usually a temperature above which they are completely miscible. For example, at room temperature a saturated solution of phenol in water contains about 8 per cent of phenol, whereas at a temperature of 68.4° C., or above, the two liquids mix in all proportions.

A few cases are known in which the solubility coefficient *diminishes* with rise in temperature, e.g., if a saturated solution of ether and water, prepared at room temperature, be gently warmed, it will become turbid, indicating partial separation of the two liquids.

6. Solutions of Solids in Liquids.—When ever a solid is brought in contact with a liquid, some of the solid gradually disappears (dissolves) and a homogeneous system is eventually obtained. This phenomenon is perfectly general, for by careful analysis it has been found that all of the so-called "insoluble substances" are in reality substances with a very low solubility. Consequently, we should call such substances "sparingly-soluble substances." In a word, then, the solubility of a solid in a liquid is always limited, and depends on the nature of the solvent and the solute (in this case the solid) and on the temperature. As in the other types of solution, we are really dealing with a condition of equilibrium, in this case between the dissolved substance and the solid. Since the degree of solubility varies greatly, some substances being so slightly soluble that the most delicate of physico-chemical methods are necessary to show that they have dissolved at all, it is evident that we can get a better idea of this class of solutions by studying the more soluble substances, such as sugar, urea, or the various salts.

When a solid like sugar is brought into contact with a solvent like water, the solid particles gradually disappear, or diffuse into the solvent in a manner quite similar to the diffusion of a gas in a closed container, for it can be shown by analysis that the dissolved particles are *equally* distributed throughout the solvent, just as any portion of a gas under the same temperature and pressure can be shown to contain the same number of particles. In this case (solution of solids in a liquid), two

forces are at work: first, the kinetic energy of the molecules, which would tend to cause them to fly off into space; second, the attraction of the solvent molecules for those of the solute. The latter force is undoubtedly the more powerful of the two, especially in dilute solutions. Hence, there is a tendency for the solvent to "pull" the surface molecules of the solid away from it. However, some of the molecules that are moving hither and thither in the solvent must return to the surface of the undissolved solid particles and be held by them. Now, as the number of particles in solution increases in number, the chances of their colliding with the solid particles will likewise increase, until finally a condition will obtain in which the number of dissolved particles returning and remaining will just equal in number those going into solution. At this stage no further solution will take place, and the mixture, being in a state of equilibrium, is said to be "saturated" for this temperature.

The weight of the solid dissolved by 100 c.c. of the solvent so as to make a saturated solution at any given temperature is the usual standard of solubility of the solid.

Influence of Temperature on Solubility of Solids in Liquids.—If fresh portions of finely ground potassium nitrate are agitated with water at room temperature (20° C.) till the water is unable to dissolve any more, it is found that 100 grams of water have dissolved about 32 grams of potassium nitrate. Such a solution is called a *cold saturated solution* of potassium nitrate.

If the potassium nitrate is added to boiling water, and agitated (stirred with a glass rod) until the water (kept at 100° C. by standing the vessel in a bath of boiling water) ceases to dissolve any more, 100 grams of water are found to have dissolved about 245 grams. This is a *hot saturated solution.*

The solubility of most solids in water increases fairly rapidly with the temperature, as in the case of the potassium nitrate just cited; but some, notably sodium chloride, are only slightly more soluble in hot than in cold water. Calcium hydroxide and certain other calcium salts are less soluble in hot than in cold water; a solution of calcium hydroxide, saturated at room temperature, becomes turbid on heating, which indicates a lower degree of solubility at the higher temperature.

Methods of Preparing Saturated Solutions.—There are two general methods employed in the preparation of saturated solutions: (1) By agitating an excess of the finely divided solute in a known quantity of

the solvent, at a definite temperature, until equilibrium is attained; and (2) by heating the solvent with an excess of the solute to a temperature higher than that at which saturation is required, and then cooling in contact with the excess of solid solute to the desired temperature. Both of these methods are satisfactory, provided sufficient time is allowed for equilibrium to be established, and provided the solute is always in excess. The first method is likely to give results that are a little too low, especially if the solid is not in a finely divided state, since the solubility of a solid depends somewhat upon the state of division. The more finely divided the solid, the more quickly is a saturated condition obtained. In fact, if a solid solute is in the form of coarse particles, a condition of equilibrium between the solid and the solution will be established at a lower degree of concentration than that at which equilibrium is established if the particles are in a state of fine division. If fine particles and coarse particles are present in the same solution, the solution may become saturated in relation to the layer of coarse particles before it is saturated in relation to the smaller particles. This is proved by the fact that the larger particles "grow" at the expense of the smaller ones.

While the results obtained by the first method are often a little too low, owing to the incomplete saturation of the solution, those obtained by the second method are generally a little too high, unless there is a considerable excess of the solute at the higher temperature and the mixture is agitated occasionally while it is being cooled to the desired temperature. In other words, if a clear portion of the saturated solution at the higher temperature is set aside in a smooth vessel, the excess of dissolved solid may not separate. Such a solution is said to be *supersaturated*.

7. Solid Solutions.—When solutions are mentioned one ordinarily thinks only of liquid solutions, i.e., solutions in which the solvent is a liquid; but it is now a well-known fact that substances dissolve not only in liquids, but also in gases and in solids. Gaseous solutions have already been described in Arts. 3 and 4 of this chapter. "Solid solutions" show all the established characteristics of other solutions, such as diffusion, saturation, etc. Just as in the case of liquid solutions, there are three types of solid solutions, namely, gases in solids, liquids in solids, and solids in solids.

(a) *Solutions of Gases in Solids.*—Many solids have the property of dissolving large quantities of gases. For example, charcoal dissolves

large volumes of carbon dioxide, and hydrogen dissolves in metallic palladium until at saturation the metal contains about 1000 times its own volume of gas. This phenomenon, known as *occlusion*, is but little understood. Van't Hoff has suggested that, in the case of solution between palladium and hydrogen, we are dealing with mixed solid solutions, i.e., a solution of hydrogen in palladium, and a solution of palladium in solid hydrogen.

(*b*) *Solutions of Liquids in Solids.*—Solids also have the property of taking up many liquids in greater or less quantities. Many of the substances that we are accustomed to think of as solids are really mixtures of solids and water. Moreover, it is very difficult to obtain these solids free from water. Such substances, which include many biological substances, must be "dried" at a fairly high temperature for a long time in order to ascertain the actual amount of solid material present.

Our knowledge of the properties of liquids in solids is very limited, owing largely to the difficulties involved in studying them. In fact, at present it is difficult to say just when chemical action between the two ceases, and true solution begins.

(*c*) *Solutions of Solids in Solids.*—When we come to solid solutions proper, our knowledge is much more satisfactory than what we have concerning either of the cases just considered. This knowledge we owe largely to van't Hoff, who published his well-known paper on the subject in 1890. The subject is not of sufficient *biological* importance to warrant discussion here. Those interested are advised to consult the original paper.[2]

[2] Zeit. phys. Chem., **5**, 322, 1890.

CHAPTER VI

WATER: THE GREATEST SOLVENT

1. Historical Sketch.—From the earliest times, water has been recognized as the greatest and most general solvent. In fact, it has long been considered one of the most important substances in the world. Ancient Greek philosophers regarded water as "the beginning of all things." Thales wrote, "All things have their origin in water and return into the same." Hippocrates is regarded as the father of hydrotherapy, as well as of medicine. He and his followers refer again and again to the healing power of water. In the seventeenth century we again find many scientists referring to the importance of water. Basil Valentine, like the Greeks, considered it the source of all substances. He says, "Water is the mother of all metals." Robert Boyle, who lived at the same time, seems to have held the same idea. To quote from his book, "Sceptical Chymist": "It seems evident that water may be transmuted into all the other elements." Again, he writes, "Not only plants, but animals and minerals may be produced out of water."

While the ideas of the alchemists were, of course, erroneous, still it is evident that their conception of water as a solvent was correct. In other words, we recognize to-day, as did the alchemists, that water is the most general and most important solvent known to man. Aqueous solutions are by far the most important in every branch of science. In chemistry, by far the greatest number of reactions take place in aqueous solution. In the field of geology, we likewise find that water is the most important agency in effecting geological changes. It is also common knowledge that aqueous solutions are fundamental for all biological phenomena. Water is the *sine qua non* of life.

If water were an uncommon liquid, there is no doubt that the masses would be impressed with its wonderful properties; however, since it is so common, we are apt to think but little of it, yet, as we shall see, it of all substances is endowed with a most extraordinary combination of properties, all of which play a part in contributing to its importance in vital processes.

114

2. Specific Heat Capacity of Water.—Of all solids and liquids under ordinary conditions of temperature and pressure, water has the highest heat capacity, or specific heat. That is to say, it takes a greater quantity of heat to raise the temperature of a gram of water one degree than is required to raise the temperature of a gram of any other substance one degree. The only liquid that exceeds water in specific heat capacity is liquefied ammonia; however, ammonia is not a liquid, but a gas under ordinary conditions of temperature and pressure. Consequently, water is taken as the unit, and the specific heats of other substances are compared with the specific heat of water.

The high specific heat of water is of value to living organisms, since the average cell contains about 80 per cent water. When heat is liberated in metabolic reactions, the temperature of the cell does not rise very greatly, for water plays the rôle of a heat regulator, that is, it takes up the heat which is liberated and gives it off gradually. If water had a low specific heat capacity, some parts of the body would get too hot and others too cold.

On a larger scale, this property of water also plays a very important part in life processes. The author refers to the fact that nearly three-fourths of the earth's surface is covered with water, and thus vast quantities of heat are transported from one part of the world to another, by the ocean currents, etc. The hot places are thereby prevented from getting excessively hot, and the colder places benefit in consequence. Plants and animals of benefit to man are thereby enabled to live in most parts of the earth.

3. Surface Tension of Water.—The surface tension value of water is greater than that of any other liquid except mercury. The relative value, as compared with some of the more common organic solvents, can be seen from table VI:

TABLE VI

SURFACE TENSION OF LIQUIDS AT 20° C.

	In dynes per cm.
Water	72.7
Glycerol	65.2
Benzene	27.9
Ethyl acetate	23.3
Ethyl alcohol	21.7
Ethyl ether	16.5

The general concept of surface tension will be considered later on, after we have considered the general properties of colloidal solutions. Briefly speaking, there is a kind of energy which manifests itself at the surface of every liquid. This energy is due to the unequal attraction between the molecules of the liquid at its surface and those within the body of the liquid. The liquid behaves as if it were surrounded by a very thin and tightly stretched skin.

In colloidal solutions (which are mostly aqueous) there is also a surface tension manifested at the surface of the different interfaces of the suspended particles. This principle has an important bearing on the phenomenon of "adsorption," which plays a most important part in physiological phenomena.

4. Dielectric Constant of Water.—Water has the highest dielectric constant, or specific inductive capacity for electricity, of any of the more common liquids. The significance of this statement lies in the fact that there is a relation between the power of a liquid to break down molecules into ions and the dielectric constant of the liquid. The relative position of water as compared with other liquids can be observed by glancing at table VII:

TABLE VII

Solvent	Dielectric Constant	Dissociating Power
Hydrocyanic acid	95	greater than water
Hydrogen dioxide......	93	greater than water
Formamide...........	84	greater than water
Water...............	82	1
Formic acid......... .	62	less than water
Ethyl alcohol..	26	about $\frac{1}{3}$ that of water
Acetone.......	21	about $\frac{1}{4}$ that of water
Ethyl ether..........	4.5	about $\frac{1}{18}$ that of water
Benzene...	2.4	very small

It will be observed that the only liquids that have a greater dielectric constant than water are unusual solvents.

Just what is meant by the term "dielectric constant" of a medium? In 1837, Faraday made the important discovery that the attraction or repulsion between two electrically charged particles depends not only

on the size of the charges and their distance from each other, but also on the nature of the medium separating the two charged bodies. In a word, then, the *dielectric constant* is a measure of the insulating capacity of the medium. Air is taken as the standard dielectric, and all other dielectric constants are referred to that of air as unity. Water possesses 82 times as much power to keep oppositely charged particles apart as does air.

It is a remarkable fact that so many of the special properties of water seem to be just the ones that are best for cell activities. The property of high dielectric constant is no exception to the rule. This property accounts for the fact that substances dissolved in water interact more readily than when dissolved in any other common liquid. The living cell contains about 80 per cent water, hence all kinds of chemical reactions are accelerated within it.

5. Constitution of Water.—Let us now inquire into the molecular structure of water. In the first place, it is made up of two of the most important elements, perhaps the two most important—oxygen and hydrogen. The usual manner of writing the formula is H_2O, since that represents the mathematical relationship of its constituent elements. The formula H_2O, however, only tells a part of the story, for on closer inspection it is found that water is really made up of the constituent parts of all acids and bases. Acids are substances that on solution dissociate a hydrogen atom carrying a positive charge of electricity, i.e., an H· ion, or, in more modern terms, a pure positive particle of electricity—*a proton*. Bases are substances that on solution dissociate a hydroxyl group carrying a negative charge of electricity, i.e., an OH′ ion. In other words, this OH′ ion carries an extra electron which it has appropriated from the hydrogen atom. Now, when these oppositely charged ions come together, we get neutral water, H·OH. Moreover, when this reaction takes place, we have the most exothermic or heat-forming chemical reaction known. The resulting compound, water, is the most stable chemical compound, just as one might expect. It is a general rule that those substances that unite with the greatest amount of heat are the most stable.

It is a noteworthy fact that most biochemical reactions involve the liberation or reunion of ions of water; nevertheless, in living organisms this reaction takes place without thermal changes. This remarkable feat is brought about through the assistance of biological catalysts or enzymes. (See Chapter XII.)

Water a Complex Mixture.—We have seen that as our knowledge of the properties of water increased we changed the formula from H_2O to H'OH. Now, while this formula is apparently the correct one for water in the *dry* gaseous state, modern investigators have come to the conclusion that water in the liquid and solid states is much more complex. In other words, the triatomic formula, H'OH, does not account for all the properties of water. In the first place, the freezing- and boiling-points are not at all what we should expect. Water is composed of gases that have extremely low freezing- and boiling-points. Oxygen boils at $-181°$ C., and hydrogen at $-253°$ C. From comparison with compounds of known composition, ice should form at $-150°$ C., and the boiling-point of water should be $-100°$ C. It appears, then, that the molecular weight of liquid and solid water must be greater than 18; that is to say, it must be a polymerized or associated molecule, in which a number of molecules are united, and the formula should be written $(H'OH)_n$. There is some difference of opinion as to the numerical value of n, but it is usually assumed that three complexes exist, in which $n = 1, 2,$ and 3, and these three kinds of molecules are represented by three formulæ: H'OH, which is called "hydrol"; $(H'OH)_2$, called "dihydrol"; and $(H'OH)_3$, called "trihydrol."

Liquid water is considered to be a *solution* composed of all three of these substances, all in chemical equilibrium with one another. Solid water, or ice, is assumed to be nearly pure trihydrol. Even water vapor, or steam, is not considered to be pure H'OH. The vapor density of steam is rather too great for this formula, just as it is too small for $(H'OH)_2$. It is therefore assumed that steam is a mixture consisting mostly of H'OH molecules with some $(H'OH)_2$ molecules. The equilibrium conditions for steam at $100°$ C. correspond to about 91 per cent $H \cdot OH$ and 9 per cent $(HO'H)_2$. Absolutely *dry* steam would probably be pure H'OH.

In accordance with the above data, a solution whose constituents are a *pure* substance, such as sugar and the associated liquid, water, has the following molecular components: $(H'OH)$, $(H'OH)_2$, $(H'OH)_3$, $C_{12}H_{22}O_{11}$, $C_{12}H_{22}O_{11} \cdot (H'OH)_n$, and possibly others. These molecules are all in perfect equilibrium with one another as long as the water is in the liquid state.

Many of the peculiar properties of water are explained quite satisfactorily by assuming it to be a polymerized compound. The large

amount of heat required to convert ice into water, and liquid water into steam, is hereby explained. As we have learned, ice is mostly trihydrol, liquid water mostly dihydrol, and water vapor mostly hydrol; therefore, the application of heat must do three things, namely, break up the polymers, heat the polymers, and heat the single molecules. The peculiar factors associated with the specific heat of water are thus explained.

Likewise, the point of maximum density of water at 4° C. can be explained on the assumption that ice, or trihydrol, exists in dihydrol, or liquid water, since ice at 0° C. has a lower density than water at 0° C.; and as water is heated from 0° upwards, there are two opposite forces at work—dilatation of the molecules, according to rule, and contraction, due to the conversion of trihydrol into dihydrol. Since the latter process predominates over the former at the lower temperatures, and the reverse is true at the higher temperatures, there must be a point where the difference between them is least.

The fact that the viscosity and compressibility of water increase to an unusual degree as the temperature is lowered is also explained on the associated-liquid theory.

There is another interesting fact in connection with water which throws some light on its properties. Water, of all known liquids (except fused metals), contains the largest number of molecules per unit volume. For instance, there are five times as many gram-molecules per cubic centimeter in water as there are in benzene. This means that there is less space between the molecules of water than between those of other liquids. This accounts for its low compressibility and explains why ice is lighter than liquid water. The molecules of trihydrol which constitute ice are not so *closely packed* as the mixed molecules which constitute liquid water.

From the foregoing statements regarding water vapor, liquid water, and ice, we must conclude that *hydrol, dihydrol*, and *trihydrol* are three distinct chemical individuals, which are readily converted into one another through very slight changes in conditions. Liquid water is a mixture of all three; but when the temperature is near the boiling-point the percentage of trihydrol is very small and the percentage of hydrol is relatively high; the reverse is true when the temperature is near the freezing-point.

The study of hydrol, dihydrol, and trihydrol has given confirmative proof of the quadrivalence of oxygen, since the correlation of all the

properties of liquid water leaves little doubt that it contains the following molecules:

$$H \cdot OH; \quad H-O-H; \quad H-O=O-H; \quad H-\overset{|}{O}-\overset{|}{O}-H; \qquad H_2$$
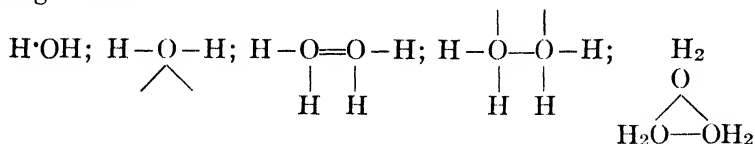
Armstrong assumes, further, that liquid water contains an isomeric form of dihydrol in which one of the oxygen atoms is quadrivalent and the other is divalent. This form is very active chemically and has the following formula:

$$
\begin{array}{ccc}
H & & H \\
 & \diagdown \diagup & \\
 & O & \\
 & \diagup \diagdown & \\
H & & OH
\end{array}
$$
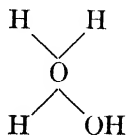
Whether water actually does contain molecules whose oxygen atoms possess extra valences, is not as yet an absolutely proved fact, but the evidence certainly points in this direction. The extra valences of oxygen would account for such properties of water as its association coefficient, its power of solution, and its ability to hydrate ions.

6. Water as a Solvent.—There is little doubt that the most important property of water, from all standpoints, is its *power of solution*, in which it surpasses all other substances. All kinds of substances dissolve in it: non-electrolytes, electrolytes, colloids, etc. The urine, for instance, may have as many as a hundred substances dissolved in it. Among these are urea, carbamic acid, creatine, creatinine, uric acid, xanthine, guanine, hypoxanthine, adinine, allantoine, hippuric acid, phenaceturic acid, oxalic acid, benzoic acid, phenolsulphuric acid, indoxylsulphuric acid, paraoxyphenylacetic acid, urobilin, urochrome, uroerythrin, hæmatoporphyrin, glucose, levulose, lactose (when the mammary gland is active), glycuronic acid, glycine, alanine, leucine, tyrocine, enzymes, putrescine, cadaverine, chlorides, bromides, phosphates, sulphates; salts of sodium, potassium, calcium, ammonia, magnesia, iron, carbonic acid, nitrogen, argon; albumins, globulins, proteoses, peptones, nucleoproteins, fibrin, oxyhæmoglobin, blood-form elements and pigments, pus, bile pigments, bile acids, acetone, acetoacetic acid, beta-hydroxybutyric acid, conjugate glycuronates, pentoses, galactose, arsenic, mercury, inositol, melanin, urorosin,

nephroroscin, urocromogen, cystine, and homogentisic acid. Only a few of these are soluble in any other one liquid.

7. Water as a Catalyst.—The whole subject of " catalysis " will be considered in a later chapter, but it seems advisable at this point to draw attention to the extraordinary rôle played by water in the "catalyzed" chemical processes of life. Most of the reactions occurring in living organisms are those in which the ions of water are involved in the process. In a large number of these reactions, the ions of water are removed or added—this phenomenon is known as hydrolysis; in others the H-ions may be replaced by OH-ions—this is the usual manner in which biological oxidations take place; in still others the OH-ions of the substance may be substituted by H-ions—this is known as reduction. All of the chemical reactions cited belong to the class known as *reversible reactions*. These reactions do not take place instantaneously; in fact, they are likely to take place very slowly unless they are assisted by some outside agency. This is where the catalyst comes in. A catalyst is a substance that changes the rate of speed of an existing chemical reaction but does not form a part of the system in its final equilibrium. Biochemical catalysts are universally present in the living cells. They bring about the chemical reactions of life processes by entering into an adsorption complex with the substance upon which they act; then, into a loose chemical union with the substrate, as is proved by the "specificity" of their action; finally, the enzyme-substrate complex attracts the ions of water. If both the H-ions and the OH-ions tend to combine with the enzyme-substrate complex, thereby resulting in the hastening of the splitting up of the substrate into simpler constituents, the process is known as *hydrolysis*. If OH-ions are substituted for H-ions on the products of the substrate, the process is known as *oxidation* (in this case an "outside" *hydrogen acceptor* must be present to take up the liberated H-ions). If OH-ions of the substrate are substituted for H-ions on the products of the substrate, the process is known as *reduction*.

These are by no means all of the extraordinary properties of water. Others, such as the peculiar nature of aqueous colloidal solutions, will be considered in a subsequent chapter.

CHAPTER VII

DIFFUSION AND OSMOTIC PRESSURE

1. Introductory Remarks. —In a previous chapter reference was made to the importance of solution in all branches of science, and special reference was made to its importance in connection with the animal and vegetable world. It has been pointed out that the supply of energy necessary for all the changes which manifest themselves in living matter arises from chemical reactions occurring in solution. (Of course, the energy originally comes from sunlight.)

Protoplasm has been referred to as a complex aqueous solution, consisting in part of substances in colloidal suspension, and in part of crystalline substances in true solution.

The importance of solution in all metabolic processes has also been indicated. No substance can obtain entrance into a living organism unless it is soluble. Most of our foodstuffs are ingested in the colloidal state and must be changed into the crystalloidal state during the process of digestion. It is only after this change has taken place that *true solution* occurs. Now, crystalloidal particles in solution have more energy than colloidal particles in solution. For example, colloidal particles in solution have little power of diffusion and are not able to obtain entrance to the living cell, whereas crystalloidal particles in solution have a remarkable power of diffusion, and therefore can pass in or out of the cells with perfect ease.

2. Osmotic Pressure.—When any substance is dissolved in a solvent, the solution, as compared with the solvent, behaves as if the solute exerted a pressure. This is evident, since the solute, after introduction into the solvent, soon diffuses throughout the entire mass of the latter. Furthermore, if one hinders this process of diffusion by first separating a portion of the solvent with a suitable partition, such as parchment paper, a pressure is developed; and, since this pressure tends to distend or push the obstructing partition, it is called "osmotic," from the Greek ὠσμός, meaning "a push."

The process of "osmosis" was first observed by Abbé Nollet, who showed that if a solution of sugar be placed in a vessel which is closed by an animal membrane and dipped in a vessel of water, the water will diffuse through the membrane and cause the solution to rise in the containing vessel. This is the phenomenon of endosmosis. The height to which the liquid rises, however, cannot be taken as a measure of the osmotic pressure of the solution, for the level of the liquid never remains stationary, but after reaching a certain height begins to fall again, until the levels outside and inside the vessel are the same. In other words, while it is a simple matter to demonstrate the existence of osmotic pressure, it is not so easy to measure it. To accomplish this feat it is necessary to have a membrane for the endosmometer which will allow the solvent to pass freely through, and at the same time prevent the passage of the solute. Most animal and vegetable membranes only partially accomplish this. Many years ago, Graham showed that such membranes, e.g., pigs' bladders and parchment paper, will allow water and dissolved crystalloids to pass, but will prevent the passage of colloidal substances, such as proteins or starch. Membranes that have the property of allowing the solvent to pass through them, and of preventing some or all of the dissolved substances from passing, are known as *semi-permeable membranes*. Animal and vegetable membranes are, therefore, said to be semi-permeable to colloids but not to crystalloids. The determination of the osmotic pressure of colloids is of relatively little value, since these substances exert but little osmotic pressure. It is evident, therefore, that in order to measure the osmotic pressure of crystalloids, other kinds of membranes have to be procured, i.e., membranes that will completely prevent the passage of dissolved crystalloids, such as sugar, urea, salts, etc., and at the same time permit the passage of such a solvent as water.

In 1867 Mauritz Traube discovered that certain inorganic precipitates are semi-permeable with respect to the salts from which they are formed, and semi-permeable also with respect to many other dissolved crystalloids. Such a semi-permeable precipitate is, for example, copper ferrocyanide. When a dilute solution of copper sulphate is mixed with a dilute solution of potassium ferrocyanide, a brown precipitate of copper ferrocyanide is produced:

$$2CuSO_4 \times K_4Fe(CN)_6 = Cu2Fe(CN)_6 + 2K_2SO_4;$$
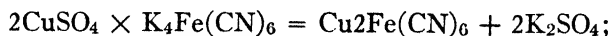
and this precipitate, while permitting the passage of water, will prevent

the passage of either copper sulphate or of potassium ferrocyanide, as well as many other crystalloids.

3. Demonstrations of Osmotic Pressure.—A good method of demonstrating osmotic pressure by utilizing Traube's discovery is as follows: Take a piece of glass tubing about 2 cm. in diameter and 10 cm. long and tie a piece of parchment paper tightly over one end. Soak this in distilled water for several hours to drive out the air bubbles. Select a *one-hole* rubber stopper that will tightly fit the tube. Pass a fine capillary tube, at least a meter long, *just* through the hole in the stopper, then fit the stopper into the *top* of the glass tube. Place the large tube in a beaker which is at least 10 cm. deep. Clamp the tube in a vertical position. Remove the stopper and carefully fill the tube with a 3 per cent solution of potassium ferrocyanide. Replace the stopper, using great care not to have any air bubbles beneath the stopper, and at the same time to prevent *forcing* any of the liquid into the capillary tube. Wash the outside of the tube with copper sulphate solution. Quickly clamp into position in the beaker, and fill the beaker with 3 per cent copper sulphate solution. Finally, set the apparatus in a quiet place for a few days, observing at repeated intervals the height to which the liquid rises in the capillary tube.

4. Pfeffer's Measurements of Osmotic Pressure.—The first satisfactory direct measurements of osmotic pressure were made by Pfeffer. Utilizing the discovery of Traube, he succeeded in preparing artificial semi-permeable membranes which satisfied all the requirements, i.e., they permitted the passage of the solvent, prevented the passage of the solute, and at the same time withstood sufficient pressure to sustain the high osmotic pressures of the solutes involved. Traube's membranes were not strong enough to use in ordinary work. Pfeffer's membranes, on the other hand, would not only withstand relatively high pressures (several atmospheres), but would yield remarkably accurate results, as has been proved by later work. An account of the apparatus used by Pfeffer and the method that he employed will be given in his own words:[1]

"I obtained the first favorable results by proceeding as follows: I took (unglazed) porcelain cells, such as are used for electric batteries, and, after suitably closing them, I first injected them carefully with water, and then placed them in a solution of copper sulphate, which, either immediately or after a short time, I introduced into the interior

[1] Scientific Memoirs Series, IV, 4–5. Edited by Prof. J. S. Ames. Amer. Book Co.

of a solution of potassium ferrocyanide. The two membrane formers now penetrate diosmotically the porcelain wall separating them, and form, where they meet, a precipitated membrane of copper ferro-cyanide. This appears, by virtue of its reddish brown color, as a very fine line in the white porcelain which remains colorless at all other places, since the membrane, once formed, prevents the substances which formed it from passing through.

"In Fig. 1 (Fig. 13 of this text) the apparatus ready for use, with manometer, *m*, for measuring the pressure, is shown, at approximately one-half the natural size.

"The porcelain cell, *z*, and the glass pieces, *v* and *t*, inserted in position, are shown in median longitudinal section. The porcelain cells which I used were, on the average, approximately 46 mm. high, were about 16 mm. internal diameter, and the walls were from $1\frac{1}{4}$ to 2 mm. thick. The narrow glass tube, *v*, called the connecting-piece, was fastened into the porcelain cell with fused sealing-wax, and the closing-piece, *t*, was set into the other end of this tube in the same manner. The shape and purpose of this are shown in the figure."

A great number of details must be carried out in order to prepare a good cell for measuring osmotic pressure, as can be seen from the following paragraphs, which are quoted from Pfeffer's monograph:

FIG. 13.

"All[2] porcelain cells were treated first with dilute potassium hydroxide, and then with dilute hydrochloric acid (about 3 per cent), and after being well washed were again completely dried before they were closed, as already described. Substances which are soluble in
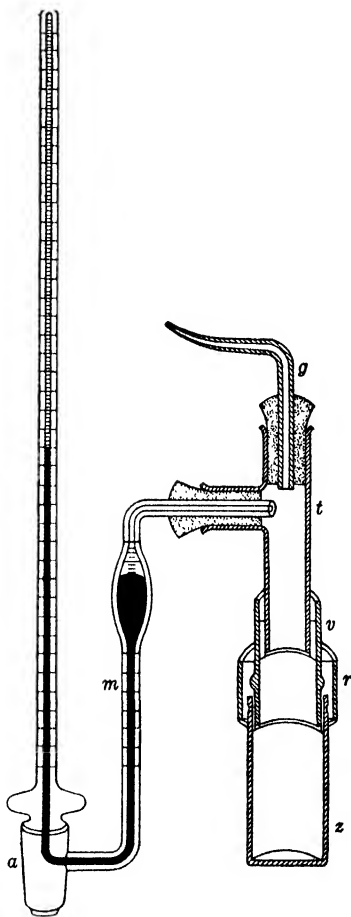
[2] Scientific Memoirs Series, IV, 6–7. Edited by Prof. J. S. Ames. Amer. Book Co.

these reagents, such as oxides and iron, which under certain conditions can do harm, would thus be removed.

"After the apparatus was closed the precipitated membrane was formed either in the wall or upon the surface, according to the principle already indicated. In order that this should be done successfully, a number of precautionary measures are necessary, and these will now be discussed. Since I experimented chiefly with membranes of copper ferrocyanide, which were deposited upon the inner surface of porcelain cells, I will fix attention especially upon this case.

"The porcelain cells were first completely injected with water under the air-pump, and then for at least some hours in a solution containing 3 per cent of copper sulphate, and the interior was also filled with this solution. The interior of the porcelain cell was rinsed out once quickly with water, well dried as quickly as possible by introducing strips of filter paper, and after the outside had dried off somewhat, it was allowed to stand some time in the air until it just felt moist. Then a 3 per cent solution of potassium ferrocyanide was poured into the cell, and this immediately reintroduced into the solution of copper sulphate.

"After the cell had stood for from 24 to 48 hours undisturbed, it was completely filled with the solution of potassium ferrocyanide and closed as shown in Fig. 13. A certain excess of pressure of the contents of the cell now gradually manifested itself, since the solution of potassium ferrocyanide had a greater osmotic pressure than the solution of copper sulphate. After another 24 to 48 hours the apparatus was again opened, and generally a solution introduced which contained 3 per cent of potassium ferrocyanide and $1\frac{1}{2}$ per cent of potassium nitrate, and which showed an excess of osmotic pressure somewhat more than three atmospheres."

Using the apparatus described above, Pfeffer performed numerous determinations of osmotic pressures of different solutions. The solution whose osmotic pressure was to be tested was placed in the "cup," which was then closed with a stopper having a manometer attached. The cup was then placed in pure water, and after equilibrium had been established the pressure was read on the manometer.

With solutions of cane sugar, he obtained the results given in table VIII:

From these results it is evident that the osmotic pressure is proportional to the concentration of the solution, since the pressure by concentration is nearly constant. The difficulties involved in performing

the experiments are responsible for the deviations from constancy, as has been proved by later work. Pfeffer also studied the influence of temperature on osmotic pressure, and showed that the pressure increases as the temperature is raised.

TABLE VIII

C = Concentration in per cent. by weight	P, Osmotic Pressure	$\dfrac{P}{C}$
1 per cent...............	53.5 cm.	53 5
2 per cent................	101 6 cm.	50 8
4 per cent................	208.2 cm.	52.0
6 per cent................	307.5 cm.	51.2

The osmotic pressures of a 1 per cent solution of cane sugar were measured at different temperatures, with the following results:

	Temperatures	Osmotic Pressures
a.	14.2°	51.0 cm.
	32.0°	54.4 cm.
b.	6.8°	50.5 cm.
	13.7°	52 5 cm.
	22.0°	54.8 cm.
c.	15 5°	52 0 cm.
	36.0°	56 7 cm.

It is evident from these results that the osmotic pressure of such a solution increases with rise in temperature. The apparent irregularities are due to the fact that different cells were used in the experiments. These are marked (*a*), (*b*), and (*c*). The results obtained in the same cell are fairly comparable with one another.

These are only a few of the classical experiments which Pfeffer performed on osmotic pressure. They did not make much stir at the time, yet they were destined to be very far-reaching in their effect on the progress of the natural sciences. His aim was to determine the magnitude of the force, and to consider its effect upon plant life. His results showed that it was a force of considerable magnitude. A 10 per cent solution of cane sugar showed an osmotic pressure of nearly 7 atmospheres. Hence, it is obvious that osmotic pressure plays a prominent rôle in plant life, for instance, in the raising of water from the roots to the top of the plants. In the case of some trees, this is

quite a performance, for an enormous amount of water is raised as much as 500 feet.

5. Applicability of the Simple Gas Laws to Solutions.—A great deal of the interest attached to Pfeffer's experiments on osmotic pressure is due to the remarkable parallelism between the properties of gases and those of dissolved substances. This parallelism was first observed and studied by van't Hoff in 1887. By a process of pure reasoning, he reached the conclusion that the osmotic pressure of a dilute solution is directly proportional (1) to the concentration of the solute, and (2) to the absolute temperature. In the first place, let us consider the relationship to Boyle's law—that the volume of a gas varies inversely as the pressure to which it is subjected, provided the temperature is kept constant—or, in the case of solutions, that the pressure is inversely proportional to the volume. To refer again to Pfeffer's results: He found that a 1 per cent solution of cane sugar gave an osmotic pressure equivalent to 53.5 cm. of mercury. If Boyle's gas law were to hold good in the case of solutions of solids in liquids, then the osmotic pressure of a 2 per cent solution should be equivalent to 107 cm. of mercury. Pfeffer found 101.6. For a 4 per cent solution the pressure should be 214 cm. Pfeffer found 208.6, and so on. It is evident then, that there is a very close agreement between the osmotic pressure values obtained by calculation by the application of Boyle's gas law and those determined experimentally by Pfeffer. That the agreement may be attributed to errors of technique, as has been proved by subsequent more accurate work. Pfeffer's apparatus was slightly faulty, and his experiments were not performed at absolutely constant termperature.

Let us now consider the relationship to the second gas law (known as Dalton's or Gay-Lussac's Law): The volume occupied by a given mass of gas, kept under constant pressure, increases as the temperature is raised, and the relative expansion is approximately the same for all gases, the increase being $\frac{1}{273}$ of the volume for each degree Centigrade rise in temperature. In the case of solutions, the analogous law would be: The osmotic pressure of a solution of a given concentration is proportional to the absolute temperature. This law is also found to hold true. On referring to Pfeffer's results in regard to the effect of increase of temperature (table on p. 127) we see that the osmotic pressure of a 1 per cent solution of cane sugar at 14.2° C. was found to be equivalent to 51 cm. of mercury. The same solution at 32° C. had an osmotic

pressure of 54.4 cm. of mercury. Now, if Gay-Lussac's Law holds true in this case, one should get 53.4 cm. (14.2° C. = 287.2° A.T., and 32° C. = 305° A.T.),

$$\frac{51 \times 305}{287.2} = 53.4 \text{ cm.}$$

The agreement, therefore, is truly significant.

The next question is: Do dissolved particles have the same energy as gaseous particles, and in consequence permit the application of Avogadro's law to solutions? Here again, van't Hoff found that Pfeffer's results yielded a satisfactory answer to the query. The law of Avogadro, applied to gases, states that equal volumes of all gases at the same temperature and pressure have the same number of particles. If the law were applied to solutions it would be stated thus: In equal volumes of solution which, at the same temperature have the same osmotic pressure, there are contained the same number of particles. This is equivalent to saying that, if Avogadro's gas law is applicable to solutions, then, *equimolecular solutions of all substances are isosmotic.*

Let us examine the facts by studying Avogadro's law and applying Pfeffer's data. The study of gases has shown that at the temperature of 0° C., one gram-molecule of gas exerts a pressure of 760 mm. of mercury when it occupies a volume of 22.4 liters, or a pressure of 22.4 atmospheres when it occupies a volume of 1 liter. The application of this principle to solutions would be expressed as follows: At a temperature of 0° C., the osmotic pressure of a solution is 760 mm. (1 atmosphere), when it contains one gram-molecule of substance dissolved in 22,400 c.c. of solution, and this same amount of substance has an osmotic pressure of 22.4 atmospheres when dissolved in 1 liter of solution.

On applying Pfeffer's data the following facts are revealed: One gram of sugar dissolved in 100 grams of water occupies a volume of 100.60 c.c., and since the molecular weight of sugar is 342, then 100.6 c.c. of this solution contain $\frac{1}{342}$ gram-molecule of sugar, and, consequently, 22.400 c.c. will contain.

$$\frac{22.400 \times 1}{100.6 \times 342}, \quad \text{or} \quad 0.652 \text{ gram-molecule.}$$

The osmotic pressure of this solution was found by Pfeffer to be 493 mm.; and, on applying the first law, we find that the osmotic

pressure of a solution containing one gram-molecule of substance dissolved in 22.400 c.c. would be

$$\frac{493 \times 1}{0.652} = 756 \text{ mm.,}$$

a value very nearly equal to that obtained with gases. Other grammolecular substances give similar results. In fact, subsequent more carefully conducted experiments gave results that were almost exactly in accord with the values obtained for gases; hence, Avogadro's gas law applies to solutions with the same degree of accuracy as Boyle's law and Gay-Lussac's law. In short, then, at any given temperature, all gram-molecular solutions of non-electrolytes have the same osmotic pressure. This is a very important law. It is not a dead dogma, but a living principle, and is of fundamental importance not only to physicists and chemists but to students of all biological subjects. It is evident, for instance, that if we know the molecular weight of a substance, we can calculate how much of it must be dissolved in a liter of solvent in order to give any desired osmotic pressure—a fact which is very important in the transfusion of any substance into the blood.

That gram-molar solutions are more practicable than percentage solutions is apparent, since no two percentage solutions are likely to give the same osmotic pressures. For example, a 10 per cent solution of glucose has only one-third the osmotic pressure of a 10 per cent solution of urea. This is due to the fact that the molecular weight of glucose in 180, whereas that of urea is 60.

6. Causes of Gas Pressure and of Osmotic Pressure.—At first sight, it is certainly a very surprising fact there should be an equality between gas pressure and osmotic pressure, for there is surely a great difference between the phenomena with which we are dealing in the two cases. In order to explain the correlation, the kinetic theory of the structure of matter has been formulated. This theory is most easily explained in the case of gases. According to the kinetic theory, gases are made up of minute, perfectly elastic particles which are ceaselessly moving about with high velocities, colliding with each other and with the walls of the containing vessel. These particles are identical (in the case of un-ionized gases) with the molecules defined by Avogadro. The volume actually occupied by the gas particles (molecules) is supposed to be much smaller than the volume filled by them under ordinary conditions, thus allowing the molecules to move about and

to be free from one another's influence except when they collide. The distance through which one molecule moves before colliding with another is known as its *mean free path*. In terms of the kinetic theory, the pressure exerted by a gas is due to the combined effect of the impacts of the moving molecules upon the walls of the containing vessel, the magnitude of the pressure being dependent upon the kinetic energy of the molecules and their number. (For the mathematical verification of this theory, consult standard texts on physics.)

Scientists are still in doubt as to the exact cause of osmotic pressure, for neither the kinetic theory nor any other theory seems to satisfy all the phenomena connected with it. Whatever views may be held as to the cause of osmotic pressure, however, the following facts are known: (*a*) What is determined in the actual measurement of osmotic pressure is the force whereby the solute attempts to enter the solvent through a semi-permeable membrane. (*b*) The osmotic pressure of a dilute solution of such a substance as sugar or urea is equal to the pressure which the sugar or urea would exert if it were in the gaseous state at the same temperature, and occupied the same volume as the solution. It should be noted, however, that this proposition is only claimed to be valid in the case of *dilute* solutions. Van't Hoff himself claimed it to be valid only in the case of "ideal" solutions, that is, solutions so dilute that the mutual action of the dissolved molecules and their actual volume, compared with the space they inhabit, are negligibly small.

It is evident from the foregoing remarks that if one were to determine the osmotic pressure of gram-molecular solutions of all substances at the same temperature, preferably at 0° C., one would find that they could be divided into three great classes, as follows:

(1) Those having an osmotic pressure of approximately 22.4 atmospheres.

(2) Those having an osmotic pressure considerably more than 22.4 atmospheres.

(3) Those having an osmotic pressure considerably less than 22.4 atmospheres.

Solutions of the first class include most of the organic compounds and are termed simple (undissociated) solutions. In this class the substances are in a molecular degree of dispersion, and hence gram-molecular solutions of this class, if Avogadro's law holds true, should and do have the same osmotic pressure.

Solutions of the second class, those whose osmotic pressure is greater than the values obtained in the first class, include the acids, bases, and salts, i.e., those substances that conduct electricity when dissolved in water, and some other solvents. Now, as these constitute a very large and important class of compounds, their failure to conform to the laws that apply to the first class prevented the general acceptance of van't Hoff's theory of solutions for a long time. The manner in which they were brought into line will be discussed in the chapter on electrolytic dissociation. (See Chapter VIII.)

There is a third class of solutions, as noted above, in which the osmotic pressure values are less than the values obtained in the first class. In this class the molecules are assumed to be associated together in clumps. Colloidal solutions belong to this class. Their osmotic pressure will be considered in the chapter on colloids. (See Chapter XI.) At this point it is only necessary to state that, since it is the *number of particles* in solution that determines the magnitude of the osmotic pressure, owing to the clumping of the particles in this class, their osmotic pressure is much less than that of the first class.

7. How Semi-permeable Membranes Act.—The consideration of osmotic pressure leads very obviously to the question: Wherein lies the efficiency of a semi-permeable membrane?, or: Why is a membrane permeable to one substance and impermeable to another? The answer to these questions has a direct bearing not only on the mechanism of osmotic pressure, but on the rôle that osmosis plays in the equilibrium between plant and animal cells and their surroundings. Many answers have been offered to the foregoing queries, but only a few of them will be considered in this text.

M. Traube, who was the first to prepare artificial semi-permeable membranes, considered that these membranes act as "atomic sieves" which allow particles of certain dimensions to pass through, but prevent the passage of particles having a relatively large mass. Traube endeavored to estimate the relative size of particles of different dissolved substances, with the help of a series of precipitation membranes. His results were far from satisfactory, but, in partial substantiation of his theory, he showed that many of the substances that showed little tendency to pass through the majority of the membranes employed had high molecular weight.

The sieve theory, however, failed to account for many other facts.

Tamman,[3] on investigating the permeability of copper ferrocyanide, zinc ferrocyanide, and gelatin-tannin, found that, out of a total of 17 dyes tested, 5 penetrated the first membrane, 7 the second, and 11 the third. If the sieve theory is correct, then the pores of the copper ferrocyanide are the smallest, and those of the gelatin-tannin the largest. But with the same dye it was found that in some cases the copper ferrocyanide membrane was more permeable than the gelatin-tannin, and in other cases the zinc ferrocyanide was more permeable than the gelatin-tannin. Such results are quite inconsistent with the sieve theory. On the other hand, Professor Abel of Johns Hopkins University (1914) found in his studies on "vividiffusion" that the rate of diffusion through collodion membranes is independent of their thickness, a fact which rather substantiates the sieve theory.

H. E. Armstrong is the author of the so-called chemical explanation of permeability of membranes. To quote his own words:[4] "The compounds which penetrate the membranes . . . are all substances which attract water presumably only to a minor extent and which exist to some extent in solution in an unhydrated condition; those which cannot penetrate it, on the other hand, probably all form hydrates of considerable stability in solution."

"I picture surfaces generally, colloid surfaces in particular, as not merely wetted by water but as more or less hydronated and hydrolated; that is to say, they are not merely wetted by water complexes but associated with hydrone, the simple fundamental molecule of which water is composed. The intra-molecular passages in a colloid membrane, if thus hydrolated, would be guarded by the attracted hydrone molecules; molecules in a solution bathing the membrane which attempted to effect an entry through such passages if hydrolated, would be seized upon and held back in virtue of the attraction which the two hydrolated surfaces—that of the membrane and that of the solute—would exercise upon one another. The hydrolated passages, however, would be indifferent to molecules which were not hydrolated —consequently such a substance as acetic acid, of which probably only a small proportion is present in solution in the hydrolated state, would gradually pass through them.

"It is obvious that the argument now put forward may be applied to the discussion of a great number of more or less obscure physiological phenomena. . . . The argument affords an explanation of the well-

[3] Zeit. phys. Chem., **10**, 255, 1892.　　[4] Proc. Roy. Soc., 1909, B. 81, 94.

known efficacy—e.g., of mercury salts and of iodine—of alkaloids as drugs. It should point the way to the production of medicaments adjusted to their purpose according as it is desired that they should penetrate this or that membrane."

At the present time the most widely accepted explanation of semipermeability of membranes is that of selective or preferential solubility. The membrane is permeable to those substances that are dissolved in it and impermeable to those substances that are insoluble in it. This view finds support in the fact that it is possible to construct osmotic cells in which adsorption by the membrane is without doubt the controlling factor. For instance, if chloroform, water, and ether are introduced into a cylinder so as to form three layers, the ether, which is soluble in both water and chloroform, will gradually pass through the water layer into the chloroform. The chloroform, however, being insoluble in water, will not pass through it into the ether in which it is soluble. The volume of the chloroform layer consequently increases in volume, or, if the movement of the water layer is prevented by means of a semi-permeable membrane, osmotic pressure is produced.

Overton, by studying living membranes, especially plant cells, also arrived at the conclusion that the efficiency of a semi-permeable membrane depends on its ability to differentiate, by solvent or absorptive power, between the substance with which it comes into contact. He states that: "The reason of the very variable permeability of living protoplasm, or better the plasma membrane for various compounds, lies in an impregnation of the membrane by bodies of similar solvent ability like the high molecular monatomic alcohols, ether, olive oil, etc. All compounds which are readily soluble in this impregnable substance pass quickly into the cell, while those substances which are considerably less soluble therein than in water pass more slowly into the cell, and indeed the slower the more the distribution coefficient changes in favor of the water. Those compounds which are practically insoluble in the impregnable substances do not pass into the cell."

There are several other theories of semi-permeable membranes which have many data to back them up, but they are not as plausible as the *preferential solubility* theory.

8. Indirect Methods Employed in the Measurement of Osmotic Pressure.—The methods thus far described measure the absolute

osmotic pressures of solutions. These methods are all long and tedious, hence other methods have been devised for measuring the relative osmotic pressures exerted by solutions of different substances.

It is clear from the data presented in the foregoing paragraphs that a solution possesses a definite amount of *free energy*, and that if we separate any given portion of a solution from the pure solvent, or from a solution of lesser concentration, by means of a semi-permeable membrane, there is a tendency to diminution of this energy. This tendency to diminution of energy is what we call osmotic pressure. It happens that this diminution can be determined experimentally by other methods than the direct ones that have been cited. For instance, it can be determined indirectly by comparison with a solution of known osmotic pressure. Solutions that have the same osmotic pressure have a number of properties in common and are said to be *isosmotic*, or *isotonic*.

Some of the indirect methods for determining osmotic pressure involve the use of cell membranes of plants or animals; others are dependent on certain physico-chemical properties that isosmotic solutions possess in common. The most important of the latter class are those depending on determinations of the *vapor pressure*, the *boiling-point*, and the *freezing-point* of solutions. The last-mentioned method, the freezing-point method, is by far the most important, since it is now used almost exclusively in biological and medical practice.

(*a*) *Plasmolytic Methods.*—Many valuable data, however, have been obtained by the employment of so-called *plasmolytic methods* which involve the use of cell membranes. The cells of plants and animals consist of protoplasm enclosed in walls which are sometimes quite strong and resistant, and hence maintain their shape and size when immersed in solutions of other substances of similar concentration. These walls are permeable to water and aqueous solutions, but they have on the inside a thin lining which is a semi-permeable membrane, and which allows water to pass through it, but is quite resistant to many dissolved substances, such as glucose and potassium malate. The Dutch botanist, de Vries,[5] was the first to make use of such cells in the determination of osmotic pressures. He used cells from the following plants: *Tradescantia discolor. Curcuma rubicaulis*, and *Begonia manicata*. The reason for choosing these plants will be seen after we have considered the principle of the method. The de Vries

[5] Zeit. phys. Chem., **2**, 415, 1888; **3**, 103, 1889.

method depends upon the shrinking or swelling of the cell sac, when the cells are immersed in a solution whose osmotic pressure differs from that of their own sap. The same phenomenon is obtained in the living cell, when it is surrounded by a solution whose osmotic pressure is greater or less than that of the cell sap. This is the phenomenon of plasmolysis. In other words, under these conditions the protoplasm is surrounded by a membrane which is semi-permeable and hence allows water to pass through but not dissolved substances; at least the passage of many dissolved substances is prevented. Consequently, by bringing the cells into solutions of varying concentrations, we can find one in which there is neither shrinking nor swelling. Such a solution is said to be *isosmotic* with the cell contents; and if we determine the osmotic pressure of this particular solution by a direct method, or by the cryoscopic method (see below), we learn the osmotic pressure of the cell contents. The method can, perhaps, be understood better if we examine Fig. 14:
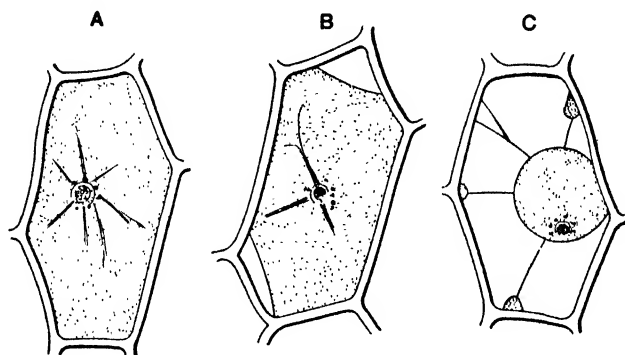

Fig. 14.

When the cell is immersed in a solution having the same osmotic pressure as its own contents, it will have the normal appearance, as shown in *A* in the figure. When the cell is immersed in a solution having a slightly smaller osmotic pressure than its own contents, it will also have the appearance of *A*, but close observation will reveal the fact that the cell is more distended than it was in the first instance. Water has passed from the solution into the cell, but the resistant cell wall has been strong enough to withstand the pressure. If, on the other hand, the cell is immersed in a solution having greater osmotic pressure than its own contents, water will pass from the cell through the mem-

brane out into the solution. The cell contents, having lost more or less water, will contract as shown in *B* or *C* in the figure; the semipermeable membrane will also contract, and this contraction can be seen quite readily, especially if the cell contents are colored. This is the reason that de Vries used the above-mentioned plants; their cells fulfilled all the necessary conditions for quick detection of any change in osmotic pressure. In other words, in order that the changes in the size of a cell may be observed, *its contents must be colored*, and all the cells must begin to show the phenomenon at the same concentration. To quote the words of de Vries, bearing on this point, "Not every plant and not every section can be employed. We need cells in which the plasmolysis can be conveniently observed, and a section in which all of the cells begin to show this phenomenon at exactly the same concentration of the external liquid."

The Procedure.—The cells are placed in a solution whose osmotic pressure is greater than their own, as shown by their contraction. This solution is continually diluted, its action on the cell at every stage of dilution being noted. At different concentrations new cells are added, and finally a solution is reached in which the newly added cells, as well as those originally present, exhibit normal form. This solution has the same osmotic pressure as the cell contents. The solution is then analyzed, and its osmotic pressure determined by one of the exact methods. Now, other solutions can be tested against these cells whose osmotic pressure is known. Any other solution will naturally be isosmotic with the cell contents, and consequently with the original one tested. In these cases it is only necessary to analyze the solution to learn what concentration is isosmotic with the original solution. Of course, dilutions that would injure the cell cannot be used.

De Vries determined the concentrations of quite a large number of substances which were isosmotic with the cell contents. He did not express the concentrations of these isosmotic solutions in terms of gram-molecules per liter of solution, but took the reciprocals of the concentrations and called them *isotonic coefficients*. These isotonic coefficients show at once the relative osmotic pressures of solutions of equal molecular concentration. A few of his results are given in Table IX.

(*b*) *The Blood Corpuscle Method.*—The principle of this method is the same as that already described with the vegetable cells.

The red blood corpuscles are an easy type of cell to work with, on account of the ease with which they can be obtained. The contents of

these cells is enclosed in a thin, elastic envelope, but, unlike the plant cells, they have no resistant cell wall to support the membrane; consequently, when they are placed in water they readily swell up and finally burst, thereby liberating the coloring matter of the cell, the hæmoglobin, so that the water acquires a deep-red color. This is called "laking."

TABLE IX

Substance	Empirical Formula	Isotonic Coefficient
Glycerol.....................	$C_3H_8O_3$	1.78
Sucrose.....................	$C_{12}H_{22}O_{11}$	1.88
Malic Acid.................	$C_4H_6O_5$	1.98
Tartaric Acid.......	$C_4H_6O_6$	2.02
Potassium Nitrate...........	KNO_3	3.00
Sodium Chloride....	$NaCl$	3.05
Potassium Acetate...........	$C_2H_3O_2K$	3.00
Magnesium Chloride.........	$MgCl_2$	4.33
Potassium Citrate...........	$C_3H_5O_7K$	5.01

Advantage of this action of red blood corpuscles was taken by Hamburger[6] for the comparison of osmotic pressures. He prepared a series of test tubes containing slightly different concentrations of solutions, and then placed a definite number of drops of defibrinated blood in each tube and allowed them to stand for a short time. In the more dilute solutions, more or less laking took place; in the more concentrated solutions, shrinking and settling to the bottom took place. There was usually one of the tubes in which no change was observable. The solution of this tube was considered to be isosmotic with the blood corpuscles.

Without going into details, it may be said that the isotonic coefficients obtained by Hamburger agree well with those obtained with the plasmolytic method of de Vries. The method, however, cannot be used for the comparison of osmotic pressures in cases in which the solution contains substances that alter the permeability of the cell envelope; for example, unsatisfactory results are obtained when the salts of ammonia, urea, alkaloids, free acids, or certain toxic substances are present. Advantage of this very fact, however, is taken to deter-

[6] Zeit. phys. Chem., 6, 319, 1890; *ibid.*, 7, 529, 1891.

mine whether or not a given substance has a damaging effect on the corpuscles, by ascertaining whether hæmolysis occurs in solutions that are known to be isosmotic with the corpuscles. Even the degree of toxicity of substances can be estimated by noting the intensity of the color liberated, as compared with some standard. There are special types of colorimeters, known as hæmoglobinometers, used for this purpose.

(*c*) *The Hæmatocrit Method.*—Solutions that cause hæmolysis, although isotonic with one another, are, of course, by no means isotonic with the fluid contents of the corpuscles, for the bursting of the cells requires a very considerable pressure. Solutions insufficiently hypotonic to cause actual rupture, as we have seen, will cause swelling through absorption of water, while slightly hypertonic solutions will cause shrinking through the withdrawal of water. It is evident, therefore, that a more satisfactory method would be one that would determine directly the concentration of solutions that are isotonic with serum, i.e., a method that would give the desired result without causing any change in the volume of the corpuscles. This is readily done by means of an instrument devised by Hedin and Koeppe, known as a *hæmatocrit*. This apparatus consists of a small graduated capillary tube, opening out at the top into a funnel-shaped expansion, to which a small syringe can be attached. Blood corpuscles, freed from serum by washing with isotonic salt solution, are suspended in measured amounts of the solutions that are being investigated, and the mixtures are then placed in hæmatocrit tubes and centrifuged for 5–7 minutes at a velocity of 8000 revolutions per minute. The relative lengths of the columns of corpuscles in the various tubes are then compared. If the corpuscles have swollen they will occupy a larger volume in the tube; if they have lost water, they will occupy a smaller volume than the corpuscles that are immersed in perfectly isotonic salt solutions.

This method has considerable clinical value, since standard saline solutions (solutions of known osmotic pressure) can be kept on hand, and the procedure with the hæmatocrit takes very little time.

(*d*) *The Cryoscopic Method of Determining Osmotic Pressure.*—The limitations in the use of the plasmolytic, hæmolytic and hæmatocrit methods in exact determinations of osmotic pressure, combined with the technical difficulties associated with the direct determination of osmotic pressure, have brought about the necessity of finding some indirect physical method that is not only easy to perform but accurate.

Fortunately, there is one physical property of solutions that can be readily and accurately measured and can be used in the estimation of osmotic pressures of solutions. This is the well-known property of solutions by virtue of which they freeze at a lower temperature than the pure solvents. When a dilute aqueous solution begins to freeze, only pure ice, free from the dissolved substance, separates out at first; then, as more and more ice separates, the solution becomes more and more concentrated, and there is a corresponding gradual drop in the freezing-point; finally, a point is reached where the solution becomes saturated, and, from that point on, ice and solute separate together. It is evident that different amounts of the dissolved substance will have varying effects on the depression of the freezing-point; that is to say, there is a direct relationship between the number of dissolved particles and the degree to which the freezing-point of the solvent is lowered. To be more specific, it has been found that when gram-molecular quantities of all non-electrolytes are dissolved in 1000 grams of water, the freezing-point is depressed 1.86° C., i.e., *all equimolecular solutions have the same freezing-point.* Now, we have already learned that all equi-molecular solutions of non-electrolytes are isosmotic; hence, if we determine the freezing-point of a solution, we can then calculate its osmotic pressure. This method is now used almost exclusively in the estimation of osmotic pressure in the type of solutions in which the physiologist is interested. Thus, the freezing-point of human blood has been found to be, on an average, −0.56° C. Since the depression of 1.86° corresponds with an osmotic pressure of 22.4 atmospheres, the osmotic pressure of the blood will be

$$\frac{22.4 \times 0.56}{1.86} = 6.74 \text{ atmospheres.}$$

The lowering of the freezing-point must be performed with great accuracy, hence a very delicate thermometer must be used. A Beckmann differential thermometer is generally used. This is a thermometer (see Fig. 15) about 60 cm. long with about 6 degrees of the scale accurately divided into hundredths of a degree Centigrade. In order to facilitate the reading, a reservoir for mercury is provided at the top of the instrument as well as at the bottom. By means of this reservoir, it is possible so to adjust the length of the column of mercury that it appears on the scale when freezing has occurred. If the temperature

is elevated, the column does not remain on the scale, in which case one must warm the mercury by means of the hand and thereby bring it back into the upper reservoir. When the freezing-point is reached while the column of mercury is below the scale, the column is brought back as before into the reservoir at the top of the tube and more mercury is brought into position. The thermometer is held in place by means of a stopper so that the mercury bulb is immersed in the fluid under examination but does not come into contact with any glass surface. A small platinum wire serves to keep the solution under examination well mixed, while a larger stirrer is used to stir the freezing mixture. (Rock salt and ice in the proportion of 1 : 3 gives a satisfactory freezing mixture.) The solution to be tested is placed in tube *A*, the thermometer and platinum stirrer are placed in position, and the tube is inserted into position in an air jacket which has previously been inserted through the metal cover of a battery jar. Both stirrers are manipulated in order to insure an equalization of temperature, and the course of the mercury column in the thermometer is observed very carefully. The mercury will *gradually fall*, and this gradual lowering of the temperature will be followed by a *sudden rise*. The point at which the mercury rests *after this sudden rise* is the *freezing-point*. This rise is due to the fact that, previous to freezing, a fluid is more or less super-cooled. As the fluid freezes, however, there is a very sudden change in its temperature, and this change is, of course, indicated by the thermometer. The freezing-point of a solution is denoted by the Greek letter Δ (delta). The Beckmann type thermometer and the Beckmann freezing-point apparatus just described are illustrated in Figs. 15 and 16.

The lowering of the freezing-point caused by the addition of gram-molecular quantities, or, to use the shorter term that is now generally used, 1 *mole* of solute to 1000 grams of solvent, is termed the "molecular lowering," or the "freezing-point depression constant." The following table gives the depression constant for a few of the more common solvents:

Acetic acid	3.90° C.	Phenol	7.40° C.
Benzene	5.00° C.	Ethyl bromide	11.90° C.
Nitrobenzene	7.07° C.	Water	1.86° C.

Note that water has a lower constant than any of the other solvents.

(*e*) *Determination of Osmotic Pressure by Rise of Boiling-point.* — The determination of the boiling-point is likewise done with a Beck-

mann apparatus; the procedure, however, is not as easy to execute as the depression of the freezing-point, hence, it is not used as often. Of course, there are times when the boiling-point method is the more practicable.

The boiling-point of a liquid is the temperature at which its vapor pressure is equal to the atmospheric pressure. In the case of a solution, provided we are dealing with a non-volatile solute, the boiling-point is higher than that of the pure solvent, owing to the fact that the solute molecules, by their attraction for the solvent molecules, keep them back from entering readily into the vapor state, so that the solution must be raised to a higher temperature before its vapor pressure equals the atmospheric pressure.
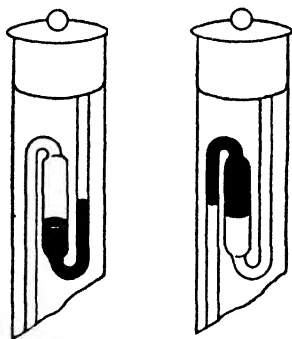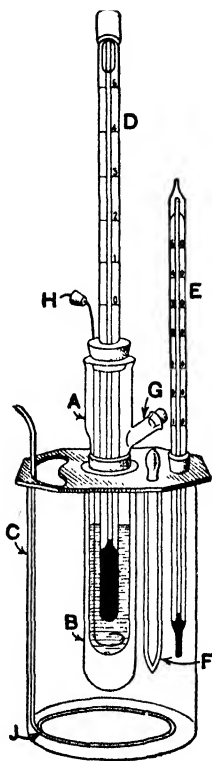

Fig. 15.   Fig. 16.

Fig. 15.—The Head of the Beckmann Type Thermometer Used in Cryoscopy, etc. The scale of this type of thermometer covers only about 6° C. and is divided into 0.01°.

Fig. 16.—Beckmann Freezing-point Apparatus. *A*, tube containing liquid to be frozen; *D*, thermometer; *H*, fine stirrer; *G*, side tube for introducing ice crystals, etc.; *B*, large outer test tube; *C*, jar containing freezing mixture; *J*, large stirrer for same.

The principle involved in the rise of boiling-point is the same as that which causes the depression of the freezing-point, that is, the rise in boiling-point of any given solvent is proportional to the number of particles of solute present in any given weight of solvent; hence, this principle can likewise be used as a means of estimating the osmotic pressure of the solution. In brief, then, gram-molecular quantities of

all solutes that are in a true molecular state of dispersion in any solvent raise the boiling-point of the solvent a definite amount. For example, the presence of 1 mole of cane sugar in 1000 grams of water raises the boiling-point 0.52° C. The figure for ethyl ether is 2.11°; for benzene 2.67°; for chloroform 3.67°.

9. Molecular Weights from Osmotic Pressure Determinations.— It is evident that, since the depression of the freezing-point and the rise in boiling-point of solutions is proportional to the number of molecules of solute present in any given weight of solvent, it is possible to determine the molecular weight of any solute. If the solute is found in the molecular degree of dispersion, it is especially easy to make the determination, but, even if the solute is an electrolyte it is possible to make the determination by the use of another factor which will be considered in the next chapter. In the case of non-electrolytes, assuming water to be the solvent, all one has to do is to find how much of the solute dissolved in 1000 grams of water is required to lower the freezing-point 1.86° C. This amount will be the gram-molecular equivalent, and from that the molecular weight can readily be calculated. It is evident that the estimation of the rise in boiling-point will give the same information, but the freezing-point method is employed oftener, since it is subject to less likelihood of error. There are cases, however, in which the boiling-point method is the more practicable.

10. Osmotic Pressure and Diffusion.— The very close connection between osmotic pressure and diffusion has already been pointed out time and again. We have seen that when a solute is placed in a suitable solvent the solute tends to become distributed throughout the entire mass of the solvent until the mixture is entirely homogeneous. Also, if two solutions of the same substance, but of different concentration, are brought together, the solute will pass from the more concentrated to the more dilute until homogeneity is established. Or, if two solutions of different substances come in contact with each other, each will distribute itself equally throughout the entire mass. In all of these cases, if the process of diffusion is interfered with by separating the parts having different concentrations with a semi-permeable membrane, then the phenomenon of osmotic pressure becomes apparent.

The general nature of diffusion of dissolved substances was very fully investigated by Graham,[7] who found, among other things, that the rate of diffusion varies greatly with the nature of the substance.

[7] Phil. Trans., **1**, 805, 1850; and *ibid.*, 483, 1851.

Substances that react quickly diffuse quickly. But the most important discovery he made was that the quantity of substance which diffuses in a given length of time is roughly proportional to the concentration of the solution originally employed.

Fick's Law of Diffusion.—It was Fick, however, who first arrived at a broad generalization in connection with the phenomenon of diffusion. He[8] stated the general law of diffusion as follows: "The dissemination of a dissolved substance through a solvent, as far as it takes place undisturbed under the influence of molecular forces alone, obeys the same laws which Fourier established for the dissemination of heat in a conductor, and which Ohm announced for the conduction of electricity. In Fourier's law it is only necessary to replace the words *quantity of heat* with the words *quantity of dissolved substance*, and the word *temperature* with *density of solution.*" In another paper he stated his law in the following terms, which will, perhaps, make the principle clearer: "The amount of salt which diffuses through a given cross-section is proportional to the difference of concentration of two cross-sections lying infinitely near to one another, or is proportional to the difference in concentration."

The law of Fick has been tested and confirmed by many subsequent workers and is now regarded as a well-established law of nature

Temperature Coefficient of Diffusion: The Principle of Soret.—If a solution is kept at a definite temperature for a period of time it will ultimately become homogeneous. The length of time, however, required to attain the state of equilibrium varies with the temperature; the colder the temperature the longer the time. If two vessels containing the same homogeneous solution are connected with a glass tube, and one vessel is kept at a higher temperature than the other, the solution will become more concentrated in the colder vessel. This phenomenon was first investigated by Soret. He established the relationship between diffusion and osmotic pressure by showing that Gay-Lussac's law also applies to diffusion. In other words, it was shown that the difference in concentration between the colder and warmer parts of the solution was $\frac{1}{273}$ of the original concentration for each degree difference in temperature. This discovery, which has been verified by van't Hoff and others, is all the more interesting since it affords another proof of the applicability of the gas laws to dilute solutions.

[8] Pogg. Ann., **94**, 65, 1865.

The Cause of Diffusion.—We now come to the question: What is the cause of diffusion? In other words, what is the force that drives the solute into all parts of the solvent until the mixture becomes homogeneous? The law of Fick establishes the fact that diffusion depends only upon difference in concentration, if the temperature remains constant. We have already learned that osmotic pressure depends upon the difference of concentration of the solution and the pure solvent. Again, we have seen that Gay-Lussac's law holds not only for osmotic pressure but for the temperature coefficient of diffusion, as evidenced by the principle of Soret. In brief, then, we can say that osmotic pressure and diffusion obey the same laws, and, this being the case, the two phenomena must either be due to the same cause, or else diffusion must be due to osmotic pressure or *vice versa*. The latter conclusion is considered to be the correct interpretation, since no common cause is known.

11. Bearing of Osmotic Pressure on Biological Phenomena.—If one undertakes a critical study of living matter one soon becomes convinced that osmotic pressure plays a very important rôle in the procedure. If osmotic pressure did not exist we would not have diffusion in solution. If we did not have diffusion, we could not have a homogeneous solution under normal static conditions. In other words, if osmotic pressure did not exist, a solution would have to be kept in a state of continuous agitation in order to be kept homogeneous. Now, living matter must be surrounded by an approximately homogeneous solution, for, as we have seen, the living cells tend to shrink, when placed in too concentrated a medium, and to swell and burst when placed in a medium that is too dilute. It is further evident that if there were no such phenomenon as osmotic pressure, and hence no diffusion, even if the living organism were placed in an approximately homogeneous solution of the right ingredients, the state of homogeneity would not be maintained very long; and, after it ceased to exist, some of the cells would tend to shrink, whereas others would tend to swell, and in time all the cells would be in such an unnatural state that they would die. Osmotic pressure also plays an important part in the supplying of food to plants and to non-mobile animals. These organisms are so constituted that their food must be brought to them in solution, and then supplied to all parts in proper proportion. It is very largely the force of osmotic pressure that accomplishes the latter phenomenon. Through this force there is an attempt to establish a homogeneous condition between the

contents of the cells of the organism and the surrounding medium, and during the process the various parts of the organism get their quota of food constituents. Of course, the condition of absolute homogeneity is never attained, since in life there is no standing still. Some of the phenomena that are responsible for this fact are purely physical. Take evaporation, for instance: Living organisms are continuously giving off moisture, which in itself is sufficient to upset the establishment of a state of absolute homogeneity.

In the higher forms of life, the multicellular organisms, other factors also play important rôles. The cells of these organisms vary in their contents according to their location; that is, the chemical composition of one group of cells of the organism may differ considerably from that of another group of cells, yet the osmotic pressure of the two groups is found to be remarkably constant. One of the other factors that exhibit themselves here is that of selective permeability of cell membranes. Other topics must be considered before this very important phenomenon is discussed; hence the detailed consideration of it will be deferred until a later chapter. (See Chapter XIII.) In brief, these membranes are capable of allowing the inflow and outflow of water and certain dissolved substances while they prevent the inflow and outflow of other substances.

Blood is one of the best examples of controlled permeability, since it is continually receiving new material from the alimentary tract and the tissues, and is continually losing water and soluble substances by various channels; yet, when one examines the osmotic pressure of this very complex fluid, one finds that this pressure is remarkably constant even if the animal has been subjected to very varied conditions. Not only is this so in the case of one animal, but the osmotic pressure of one animal does not differ much from that of another.

The cryoscopic method, described in Art 8 (*d*) of this chapter, is the most satisfactory one for determining the osmotic pressure of blood. The presence of corpuscles and colloidal suspensions and emulsoids does not have any appreciable effect on the osmotic pressure results. (The reason for this fact will be made clearer in the chapter on colloids, Chapter XI.) Hence, one can make the tests on whole blood, blood plasma, or serum, and get but slight differences in the results.

The osmotic pressure of human blood has been determined under every conceivable condition by the cryoscopic method, and has a Δ

value of about $-0.56°C.$, which corresponds with a pressure of about 6.8 atmospheres. There is, of course, some variation in a given individual. The pressure reaches a minimum at about 8 A.M., when it has a value of about $-0.52°$; it reaches a maximum pressure an hour or two after the heavy midday meal, when it has an average value of about $-0.58°$.

The osmotic pressure of the blood of other warm-blooded mammals is remarkably similar to that of human blood. The freezing-points of the blood of various mammals are as follows:

Mammal	Freezing-point Δ	Mammal	Freezing-point Δ
Man............	$-0.56°$	Sheep...........	$-0 62°$
Ox.	$-0.58°$	Pig............	$-0.62°$
Horse.........	$-0.56°$	Dog...........	$-0.57°$
Rabbit.........	$-0.59°$	Cat............	$-0.64°$

The freezing-point of mammalian blood is, hence, about $-0.592°$ C. This depression of the freezing-point would mean an osmotic pressure of $-0.592/1.86 \times 22.4$ atmospheres, or 7.2 atmospheres. In these cases, also, the hourly variations correspond quite favorably with those obtained with human blood. We see, therefore, that in mammals there is a mechanism provided for regulating the osmotic pressure of the blood, which means that the tissues are surrounded at all times with a fluid of almost constant osmotic pressure. This condition does not always obtain, however, under unhealthful conditions.

Of the various organs of the body, the kidneys seem to have most to do with the regulation of the osmotic pressure of the blood. These organs permit the outward passage of the unwanted constituents. The normal healthy kidney has the power to permit the outward passage of just the right amount of constituents to maintain the normal osmotic pressure of the blood. In case one kidney is diseased, the other kidney seems to be able to do the work of both, for the blood of individuals with only one diseased kidney does not show a very wide variation in osmotic pressure from the normal; but if both kidneys are diseased, the pressure of the blood may rise to nearly double the normal value. The urine of these individuals shows a still wider variation. The variation of the Δ-value of normal urine (24-hour specimens) is $1.30°$ to $2.30°$, whereas a variation of the Δ-value of $0.075°$ to $2.70°$ is often obtained during a 24-hour period in pathological cases.

The study of the osmotic pressure of the blood is of value in connection with many other physiological and pathological processes. For instance, in perfusion experiments, and in cases where it is necessary to introduce substances into the blood, it is necessary to employ solutions that are isosmotic with the blood. Thus, when small quantities of water-soluble drugs are introduced into human blood, they are often dissolved in a 1 per cent solution of sodium chloride, which gives about the right osmotic pressure.

The osmotic pressure of the blood of cold-blooded animals does not follow the same rule as that of warm-blooded animals. In cold-blooded animals the osmotic pressure of the fluids inside the body varies with that of the fluid in which the animal lives. The normal freezing-point of sea water is about $-2.3°$ at the surface, and the normal freezing-point of the blood of many forms of marine life is not far from this figure. Such forms of life suffer greatly if placed suddenly in water of lesser or greater concentration.

The more complex the organism, the more constant does the osmotic pressure of its blood become, and the more independent of the surrounding medium. The scale ascends from the lowest forms of marine fishes through the higher fishes to the reptiles, and finally to the highest forms of vertebrates.

One further point may be referred to before leaving the subject of osmotic pressure and its action, viz., *the poisonous action of pure water.* If tissues or cells are placed in distilled water, passage of water into the cells occurs, owing to the difference of osmotic pressure. The cells swell up and may finally burst and die. In this way pure water will act as a poison to protoplasmic tissues. A similar poisonous action on cells is observed when distilled water is drunk. In this case the surface layers of epithelium of the stomach undergo considerable swelling; salts may also pass out, and the cells may die and be cast off. This may lead to catarrh of the stomach if the practice is kept up too long. It is to this action of pure water that the harmful action of melted snow and ice is due.

CHAPTER VIII

THE NATURE AND BEHAVIOR OF ELECTROLYTES IN SOLUTION

1. Brief Historical Sketch.—A series of experiments performed during the latter half of the eighteenth century led the investigators of that period to the conclusion that electrical phenomena of different kinds are very closely related. Electricity, regardless of its source, causes definite effects according to definite laws. This was found to be true whether the electricity was produced by friction, by air currents in thunderclouds, by heat, by chemical action, or by movement of an electrical conductor in a magnetic field.

About the year 1800, Nicholson and Carlisle, of England, showed that, when a current of electricity was sent through water, hydrogen and oxygen gases were obtained at the opposite electrodes. Sir Humphry Davy, in following up the last-mentioned experiments, found that the gases that separated at the respective electrodes did not necessarily come from the same molecules. He did this by placing the electrodes in two separate beakers of water and then connecting them by placing a finger of one hand in one of the beakers and a finger of the other hand in the other beaker. The gases appeared just as before. This discovery only tended to confuse the experimenters more than ever. Davy also studied the electrolytis of aqueous solutions of acids, bases, and salts. Later, he decomposed fused sodium and potassium hydroxides by means of a current, and obtained metallic sodium and metallic potassium. In this manner he proved that there is a very close relation between the force that holds together the constituents of chemical compounds and electricity.

About this time (1805) Grotthus advanced the theory that the atoms which make up the molecules are all firmly bound together until a current is passed. In other words, he regarded chemical phenomena as the result of the phenomena of galvanic electricity. "A galvanism manifests itself in a liquid due to the heterogeneous elementary parts

contained in it. No chemical action manifests itself when the electrical forces are in equilibrium. When the equilibrium is destroyed chemical action takes place." In 1808 Grotthus showed that electrolyzed hydrogen contains positively charged hydrogen atoms and negatively charged hydrogen atoms. He concluded that, since the hydrogen molecules are neutral, they are probably composed of a positive atom and a negative atom which, when united, would give a neutral molecule. He thus predicted the modern concept of the atom nearly a century before it was experimentally established by Thomson. He also had advanced ideas on solutions: "The solution of a salt in water appears to be only an interpolation of its own elementary parts in the active galvanic molecular circle of water, as the elements of the salt contribute to the galvanic molecular activity of the water. Salts which cannot do this we call insoluble."

About 1818 Grotthus promulgated the idea that dissolved particles are continually changing partners. In a salt like potassium bromide, the potassium that at one moment is combined with any given bromine atom is combined the next moment with a different bromine atom, this exchange of parts going on continually. The relation of these conceptions to the theory of electrolytic dissociation will become apparent when we come to the consideration of that theory.

About 1832, Faraday [1] carried out his epoch-making researches on electrolysis. Much of our present-day nomenclature we owe to him. Thus, the products appearing at the electrodes and apparently moving through the solution, he called "ions." The ions that separate at the positive electrode he called "anions," meaning "to go up"; those that separate at the negative electrode, he called "cations," meaning "to go down." The corresponding electrodes he called "anode" and "cathode." The conducting substance he called an "electrolyte." It must be remembered, however, that although Faraday invented these terms, he did not have a clear conception of ionization as we now use the term; his terms apply to the products of electrolysis, not to the means of conductance.

In 1834 Faraday discovered the law which bears his name, and which enables us to predict how much chemical change will take place in an electric cell when a definite amount of electricity is passed through it. He formulated this law as follows: *The amount of salt decomposed by the electric current is proportional to the amount of electricity passed*

[1] Exp. Res. III, Ser. No. 373, 1832.

through the solution. This law is of importance since it gives us a ready means of measuring quantity of electricity. If we know the amount of decomposition of a salt like copper sulphate effected by passing, say, one coulomb of electricity, it is only necessary to pass the current we are investigating through a solution of copper sulphate and weigh the amount of copper deposited upon the cathode, in order to know how much current was passed through the solution.

Faraday's Second Law.—Faraday also found that the amounts of different metals thrown out by the same current are proportional to the atomic weights of the metals, provided all the metals in question have the same valence. If they have different valences, he found that the *amounts of the metals thrown out by the same quantity of electricity are proportional to the atomic weights of the metals divided by their respective valences.* This is the second law of Faraday.

Williamson[2] assumed an exchange of radicals in solutions of mixed salts; he made the assumption that, in a solution of hydrochloric acid, exchange occurs between the hydrogen and chlorine atoms in different molecules. To quote his own words: "We are thus led to the assumption that in a number of molecules of any compound, a continual exchange is going on between the elements in the compound. Let us take, for example, a vessel containing hydrochloric acid; it is filled with a large number of molecules of HCl. The above considerations would lead us to the assumption that an atom of hydrogen does not remain quietly attached to the atom of chlorine with which it was first combined, but there is going on a continued exchange of place with other hydrogen atoms. This change is for us, of course, not directly detectable, because one atom of hydrochloric acid is just like another."

Clausius[3] adopted Williamson's theory of interchange of radicals on kinetic grounds, but he also assumed it to take place to a very small extent only. He concluded that an electrolyte, while consisting mainly of whole molecules, contained also some part-molecules. He believed that in an electrolyte there was a constant interchange of part-molecules, whether a current was passing or not. During this interchange there must necessarily be at any given moment some free part-molecules; and these, he thought, were responsible for the conductance of the current. This assumption agreed, he claimed, with

[2] Ann. Chem., **77**, 37, 1851.
[3] Ann Physik., **101**, 338, 1857.

the fact that only a very small E.M.F. is required to electrolyze a salt. None was required to break up the molecules. In short, the only work required of the current was to direct the course of the part-molecules, which were already free and oppositely charged. The historical importance of the theory of Clausius will be seen when we come to consider the theory of electrolytic dissociation.

Between the years 1869 and 1879, Friedrich Kohlrausch[4] conducted his researches on the conductivity of acids, bases, and salts in solutions. He discovered the law, which bears his name, of the independent migration velocity of the ions. This law is: "The conductivity of any solution, referred to molecular quantities of dissolved substance—the molecular conductivity—is the sum of two constants, the one depending on the value of the cation, the other on the anion." He showed that the conductivity of different electrolytes differed greatly, and that for any given electrolyte it increased with dilution, finally reaching a maximum in most cases.

In 1885, a French chemist, F. M. Raoult,[5] published a paper on the lowering of the freezing-point and of the vapor tension of solvents by dissolved substances. This work has become a classic. The following paragraphs are a direct translation from it:

"The preceding relations show that in all the salts which it enters, every radical plays nearly the same rôle and has the same constitution; and that it always produces the same lowering of the freezing-point, which is nearly independent of the number and the nature of the radicals with which it can combine."

"The molecular lowering of the freezing-point, produced by its salts formed by strong monobasic and dibasic acids, is approximately the sum of the partial molecular lowerings of their electropositive and electronegative radicals."

"The diminution in capillary height, the increase in density, the contraction of protoplasm (de Vries' work on osmotic pressure), the lowering of the freezing-point, in brief, most of the physical effects produced by salts on the solvent water are the sum of the effects produced separately by the electropositive and electronegative radicals which are contained in them, and which act as if they were simply mixtures in the liquid."

"This shows that salts in solution in water ought to be regarded as

[4] Wied. Ann. **4**, 168, 1879.
[5] Ann. chim. phys., **4**, 104, 1885.

systems of particles, every one of which is formed by solitary atoms, and, notwithstanding its condition of combination with others, largely preserves its own individuality, its action, and its characteristic properties."

"The special constitution which we are compelled to recognize for true salts dissolved in water, does not hold for other substances, not to any appreciable extent in solvents other than water."

"The metallic salts themselves, in solvents other than water, show nothing peculiar, and produce the same molecular lowering as all other substances."

Raoult concludes his paper thus: "From this we must conclude that the subdivision of the saline molecules into electropositive and electronegative radicals takes place in fact only in aqueous solutions; and consequently, this is a result of a peculiar chemical act exerted by water on salts dissolved in it."

With all the foregoing facts in mind, van't Hoff[6] published his epoch-making paper on "The Rôle of Osmotic Pressure in the Analogy between Solutions and Gases." The details of this paper were presented in the chapter on osmotic pressure (see p. 128 *et seq.*). In this chapter, the author referred to the fact that van't Hoff and others found that it was only those compounds known as non-electrolytes that exactly obeyed the three fundamental laws of gas pressure when brought into solution, and that even these only obeyed the laws when in dilute solution. In brief, van't Hoff found that many types of solutions gave abnormal osmotic pressures, abnormal freezing-point depressions and abnormal elevations of the boiling-point; these types are: (*a*) very concentrated solutions of non-electrolytes, (*b*) colloidal solutions, and (*c*) solutions of acids, bases, and salts, i.e., true electrolytes, or, to be more specific, those substances that are capable of conducting the electric current when dissolved in certain solvents, such as water. It was this last-mentioned type of substances that gave van't Hoff the greatest concern. His own words in connection with the discrepancy which these substances showed will be given.[7]

"If we are still considering 'ideal solutions,' a class of phenomena must be dealt with which, from the now clearly demonstrated analogy between solutions and gases, are to be classed with the earlier so-called deviations from Avogadro's law. As the pressure of vapor of ammo-

[6] Zeit. phys. Chem., **1**, 481, 1887.
[7] *Ibid.*, **1**, 500.

nium chloride, for example, was too great in terms of this law, so, also, in a large number of cases, the osmotic pressure is abnormally large, and, as in the first case, as was afterwards shown, there is a breaking down into hydrochloric acid and ammonia, so also with the solutions we would naturally conjecture that in such cases a similar decomposition had taken place. Yet it must be conceded that anomalies of this kind existing in solutions are much more numerous, and appear with substances which it is difficult to assume break down in the usual way. Examples in aqueous solutions are most of the salts, the strong acids, and the strong bases. . . . It may then have appeared daring to give Avogadro's law for solutions such a prominent place, and I should not have done so had not Arrhenius pointed out to me, by letter, the probability that salts and analogous substances break down into ions."

In the final sentence we have the connecting link between the generalization reached by van't Hoff and the present-day theory of electrolytic dissociation, which was first put forth in its modern form by the Swedish physicist, Svante Arrhenius.

Arrhenius' classical paper, "On The Dissociation of Substances Dissolved in Water," appeared in the same volume of the Zeitschrift für physikalische Chemie[8] as the paper by van't Hoff, which has just been considered. Arrhenius seems to have been very much impressed with the conclusions reached by van't Hoff from his studies on solutions. The extent of his appreciation can best be shown by quoting from his paper:

"In a paper submitted to the Swedish Academy of Sciences on the 14th of October, 1885, van't Hoff proved experimentally as well as theoretically the following unusually significant generalization of Avogadro's law: The pressure which a gas exerts at a given temperature, if a definite number of molecules is contained in a definite volume, is equal to the osmotic pressure which is produced by most substances under the same conditions if they are dissolved in any given liquid.

"Van't Hoff has proved this law in a manner which scarcely leaves any doubt as to its correctness. But a difficulty which still remains to be overcome is that the law in question holds only for 'most substances,' a very considerable number of aqueous solutions investigated furnishing exception; and in the sense that they exert a much greater osmotic pressure than would be required from the law referred to.

[8] Zeit. phys. Chem., 1, 631, 1887.

"If a gas shows such a deviation from the law of Avogadro it is explained by assuming that the gas is in a state of dissociation. The conduct of chlorine, bromine, and iodine at higher temperatures is a very well-known example.

"The same expedient may, of course, be made use of to explain the exceptions to van't Hoff's law; but it has not been put forward up to the present, probably on account of the newness of the subject, the many exceptions known, and the vigorous objections which would be raised from the chemical side to such an explanation."

Other investigators, such as de Vries in his researches on plasmolysis, and Raoult in his investigation of the freezing-points of different solutions, found the same discrepancies as van't Hoff. In brief, all the methods for estimating osmotic pressure seemed to point to the fact that electrolytes exert a pressure in molar concentration which is greater than that of non-electrolytes.

The following examples will serve to illustrate the facts: If one determines the freezing-point depression of a gram-molecular solution of sodium chloride (58.5 grams in 1 liter) the amount of depression is found to be $-3.35°$ C. instead of the standard figure of $1.86°$ obtained with gram-molecular solutions of non-electrolytes. The osmotic pressure of such a solution is found to be about 40 atmospheres instead of 22.4, whether estimated directly or calculated from the freezing-point depression.

A gram-molecular solution of sodium chloride, therefore, behaves as if it contained a greater number of particles than gram-molecular solutions of non-electrolytes. The relation expressed mathematically is: $3.35/1.86 = 1.8$, i.e., it acts as if it contained 1.8 gram-molecules per liter.

In the case of other salts also, as well as most acids and bases, there is an equally pronounced discrepancy between their osmotic pressure and that of non-electrolytes.

2. The Theory of Electrolytic Dissociation.—Arrhenius, on considering the various facts that were mentioned in the foregoing pages in regard to solutions, was led to the conclusion that the anomalous osmotic pressures of solutions of acids, bases, and salts could be explained by assuming that these substances were broken down by the action of the solvent into part-molecules or ions, just as the exceptional behavior of certain gases had been completely reconciled with Avogadro's law by assuming a dissociation of the vaporized molecule

into two or more simpler molecules. The action of the sodium chloride that was considered above could be explained, he claimed, by assuming that about 80 per cent of the sodium chloride molecules were split up into the ions Na^+ and Cl^-.

Arrhenius thus went back to the theory advanced by Williamson in 1851, and expanded by Clausius in 1856; but he added to this theory a new and all-important feature, namely, a method for determining just what percentage of the whole molecules is broken down into ions, thereby converting a purely qualitative suggestion into a quantitative theory which can be tested experimentally. In other words, he showed how the degree of ionization can be calculated from conductivity and from measurements of the depression of the freezing-point. Moreover, he found that the values calculated by these two methods agree remarkably closely. This he considers the crucial test of his theory. He proposed, therefore, that aqueous solutions of acids, bases, and salts are to be considered as dissociated, to a greater or less degree, into positively and negatively charged particles or *ions*, and that the increase in the number of dissolved units which accrue from this dissociation accounts for the enhanced osmotic pressure. In other words, when such substances as hydrochloric acid, potassium hydroxide, and potassium chloride are dissolved in water, they dissociate in the following manner:

$$HCl \rightleftarrows H^+ + Cl^-,$$

$$KOH \rightleftarrows K^+ + OH^-,$$

$$KCl \rightleftarrows K^+ + Cl^-.$$
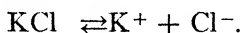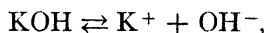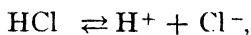
Each compound dissociates into a positively charged part called a *cation*, and a negatively charged part called an *anion*. These ions may be charged atoms like the above cations, or groups of atoms like the anion OH^-. The positive ion of an acid is always hydrogen, and the negative ion is the radical characteristic of the given acid. Thus, sulphuric acid ionizes according to the equation:

$$H_2SO_4 \rightleftarrows 2H^+ + SO_4^=.$$

The character $\rightleftarrows$ indicates that there is a chemical equilibrium between the two ions and the undissociated molecules; that is to say, the dissociation of the H_2SO_4 at finite concentrations is not complete, a certain number of undissociated molecules always being present in the solution. As the solution is diluted more and more, however, the

number of undissociated molecules continually decreases, and finally at infinite dilution the H_2SO_4 is completely dissociated. At this point there would be three times as many particles in solution as there would be if the substance were undissociated. Under these conditions the isotonic coefficient would be 3, and the lowering of the freezing-point would be $-5.58°$ C. (3×1.86).

The positive ion of a base is usually a metal; the negative ion is always hydroxyl (OH). In those cases in which the positive ion is not a metal, it is usually a positively charged group acting in the same capacity, e.g., NH_4^+; thus,

$$NH_4OH \rightleftharpoons NH_4^+ + OH^-.$$

The positive ion of a salt is the same as that of a base, and the negative ion is the same as that of an acid. For example, the salt K_2SO_4 ionizes so as to yield the same positive ion as KOH, and the same negative ion as H_2SO_4.

It may be said in passing that most of the strong acids, bases, and salts are almost completely ionized at a dilution of 0.001 mole, i.e., one gram-molecule in 1000 liters of water. We may therefore make the following generalization: The freezing-point depressions of dilute molar solutions of all acids, bases, and salts are greater than would be obtained from similar concentrations of non-electrolytes; and these depressions, as the dilution increases, approach as a limit values that are simple multiples of $1.86°$.

It was further pointed out by Arrhenius that the substances that exhibit abnormal osmotic effects when dissolved in water are the same substances that were known to conduct the electric current in aqueous solution. In other words, only *electrolytes* are capable of undergoing ionic dissociation; hence he termed the hypothesis the *electrolytic dissociation theory*.

There have, of course, been many conscientious objectors to the theory of electrolytic dissociation. It has been claimed, for instance, that since such substances as potassium or chlorine cannot remain in the presence of water under ordinary conditions without acting upon each other, it is self-evident that potassium chloride cannot break down into so-called potassium ions and chlorine ions. This objection is, of course, based upon an imperfect knowledge of the theory. No one has ever claimed that a compound like potassium chloride dissociated in the presence of water into atomic or molecular potassium and

chlorine. The products of dissociation are potassium ions and chlorine ions, i.e., positively and negatively charged particles; or, as we have it to-day, the dissociated potassium particle is a potassium atom which has lost one of its electrons, and in virtue of this loss it has not only acquired a positive charge, but its chemical properties have changed; the dissociated chlorine particle has acquired the electron (negative electrical charge) lost by the potassium particle, and in virtue of this acquisition has become negatively charged, and has become endowed with new chemical properties. Furthermore, it has relatively recently been proved that during the process of solution the ions attach to themselves a certain number of water molecules (see p. 171) and hence are able to move independently of each other. In brief, then, ions have different chemical properties from their parent atoms, they lead an independent existence in solution, and they exert the same physical effect as a molecule in solution.

3. Cause of Ionization.—The modern view of the structure of atoms helps us to understand why certain substances, when heated to high temperatures or dissolved in certain solvents, are so easily dissociated into ions. In the chapter on Atomic Structure we learned that every atom is presumed to contain a number of *external electrons* equal to its atomic number—2 for helium, 6 for carbon, 8 for oxygen, 17 for chlorine, 26 for iron, and so forth. Chemical reactions are presumed to consist in the transfer of electrons from one atom to another, or in the association of atoms in such a way as to share certain electrons.

The inert elements, the gases of Group O, owe their inertness (chemical inactivity) to the unusually stable arrangement of their external electrons. This means that they have no tendency to part with any of their electrons or to acquire any new ones.

In all other cases except those of the inert elements, there is an incomplete outer shell of electrons, and these elementary substances react so as to gain or lose electrons and thereby approach the stable arrangement of the inert gases.

The *positive* monovalent, divalent, and polyvalent elements tend to lose 1, 2, or more of their external electrons. The negative monovalent, divalent, and polyvalent elements tend to acquire 1, 2, or more electrons.

Atoms of many heavy metals, such as Cr and Mn, also tend to associate themselves with oxygen and other non-metals, by sharing electrons, to form non-metallic radicals or ions.

If we keep the facts just mentioned in mind, it is easy to see why a salt like sodium chloride is almost completely ionized when it is dissolved in such a solvent as water. Each sodium atom has one more electron and each chlorine one less electron than it needs to make a stable grouping. The sodium atom, therefore, tends to give up this electron, and the chlorine takes it up, and we have in consequence Na^+ and Cl^-. These ions have a relatively stable structure which approaches the chemical inertness of atoms of Group O; hence, they remain uncombined with each other, and are free to move off in opposite directions if an electric current is passed through the solution.

In ionizable acids we have so-called *polar combination* between the hydrogen and the acid radicals; hence, when such a compound is in solution in a solvent of high dielectric constant, there is a more or less complete separation of the hydrogen from the rest of the acid. In bases we have polar combination between the metal and the hydroxyl group.

Inside such radicals as OH^-, $SO_4^=$, NH_4^+, etc., it is evident that only a sharing of electrons occurs; otherwise, we should expect a substance like sulphuric acid to give off oxygen and sulphur ions as well as hydrogen ions.

In an acid like acetic acid, only one hydrogen is in *polar combination*; the others are combined by the sharing process with the carbon and oxygen atoms. Therefore, when acetic acid is dissolved in water, we have H^+ and $C_2H_3O_2^-$.

Furthermore, since substances like glucose or urea do not ionize at all, we must assume that in these cases all the atoms are combined by the sharing process.

The modern theory of atomic structure also explains why certain salts ionize freely, whereas the acids or bases from which these salts are derived do not. Hydrogen consists of one proton and one electron. The bond between them is very strong, hence there is little tendency to ionization. The hydrogen much prefers to share its lone electron with some other atom. A metal like sodium, on the other hand, will readily part with the electron of its outer shell if opportunity offers. Consequently, sodium acetate readily ionizes, while acetic acid does not.

To summarize: Atoms of metals tend to lose electrons, thereby forming positive ions; and atoms of non-metals tend to gain electrons, thereby forming negative ions. The ions that are thereby formed are relatively stable.

The atoms in radicals and in non-ionizable substances are combined by the sharing process.

4. Electrical Charges Carried by Ions.—If we grant the existence of ions, i.e., electrically charged fragments of molecules, we are naturally confronted with the query: Do all ions carry the same electric charge, or do some carry more than others? In the foregoing discussion, evidence has been advanced to prove that the charge carried by an ion is due to the presence of one or more electrons more or less than the number necessary to make a neutral atom. The actual value of an electronic charge has been found to be 4.774×10^{-10} electrostatic units; hence, the magnitude of the charge of a monovalent positive ion or a monovalent negative ion should be equal to the above figure; the electrical charge of a divalent or trivalent ion should be twice or three times this figure. Innumerable conductivity experiments have proved that such is the case, thereby proving that each individual ion carries an electrical charge directly proportional to the valency of the corresponding atom or radical.

5. Dissociating Power of Solvents and Their Dielectric Constants.—Water is not the only solvent that produces conductivity solutions by causing dissociation of the dissolved substances into ions. Other solvents, such as ethyl alcohol, methyl alcohol, acetic acid, formic acid, liquid SO_2, liquid HCN, liquid NH_3 and liquid HF, form solutions of fairly high conductivity with salts and other substances; these liquids are consequently termed *dissociating solvents*. Solutions in some solvents, such as ether, chloroform, and benzene, are practically non-conductors. In these solvents, typical electrolytes are often present in the form of molecules, and hence give normal osmotic pressures. In other cases the molecules are present in aggregates, hence they give in these instances subnormal osmotic pressures. These solvents are, therefore, often termed *associating solvents*. It is not known, however, whether they cause the association or simply fail to effect the breaking down of the natural polymerized solutes into molecules.

In the chapter on water (p. 116 *et seq.*) the author called attention to the fact that water has the highest dielectric constant, or specific conductive capacity of electricity, of any of the more common solvents. Knowledge of this fact led J. J. Thomson[9] and Nernst[10] to suggest

---

[9] Phil. Mag., **5**, 36, 320, 1893.
[10] Zeit. phys. Chem., **13**, 531, 1893.

that the ionizing of all solvents is in some way connected with their dielectric constants. Many investigations have proved this surmise to be correct. One of the best relatively recent investigations along this line is that of Walden.[11] Among other facts, he found that in the case of solutions of one and the same electrolyte in various solvents which have the same degree of ionization, the product of the dielectric constant and the cube root of the dilution for all such solutions has practically a constant value. Thus, if $\alpha = \alpha_1$, then $\epsilon \sqrt[3]{V} = \epsilon_1 \sqrt[3]{V_1}$; where α = the degree of ionization, ϵ = the dielectric constant, and V = the dilution.

Before continuing with the discussion it may be well to reiterate what is meant by the term "dielectric constant." In 1837 Faraday made the discovery that the attraction or repulsion between two electrically charged particles depends not only on the size of the charges and their distance from each other, but also on the nature of the medium separating the two charged bodies. In short, then, the *dielectric constant* is a measure of the insulating capacity of the medium.

Why should there be any relation between the dielectric constants of media and their own power to dissociate molecules dissolved in them? Jones[12] has answered this query in the following manner:

"Take for example, a solvent like water, which has a large dielectric constant. The force of attraction between the positively charged cation, and the negatively charged anion, of the electrolyte dissolved in water, is inversely proportional to the dielectric constant of the solvent. This being large in the case of water, the force of attraction between the positively and negatively charged parts is small. There being but a small force to hold the cation and the anion together in the molecule they separate or undergo electrolytic dissociation readily.

"If the dielectric constant of the medium is small, as in the case of hydrocarbons, the force of attraction between the oppositely charged parts is large, and they do not separate or undergo electrolytic dissociation to any appreciable extent."

6. Degree of Ionization.—By " degree " of ionization is meant the fraction ionized. Except in very dilute solutions, only a fraction of the molecules of electrolytes is dissociated into ions. Before we can predict the osmotic pressure and certain other properties of electrolytic solutions, it is necessary to determine the degree of ionization, i.e., the

[11] Zeit. phys. Chem., **54**, 133, 1906.
[12] Nature of Solutions, D. Van Nostrand Co., New York p. 210.

fraction of the solute which is ionized. This can be determined by several methods, two of which will be outlined.

The method that is most commonly used to determine the degree of ionization is the cryoscopic method. For a substance such as sodium chloride, one that dissociates into two ions, the lowering of the freezing-point at great dilution approaches, but never exceeds, twice that obtained with a solution containing the same fraction of a gram-molecule of a non-electrolyte in a liter. A solution of sodium phosphate, Na_3PO_4, which can furnish 4 ions, produces a lowering of the freezing-point, at great dilution, which approaches, but never exceeds, four times the normal. We know that in those cases in which there is no ionization, the lowering of the freezing-point in water, when 1 mole (gram-molecule) is dissolved in 1000 grams of solvent, is $1.86°$ C. Now, since the depression of the freezing-point depends on the total concentration of solute particles, in solutions of electrolytes it will depend on the concentration of un-ionized molecules plus the ionized particles; i.e., the lowering of the freezing-point will be greater than $1.86°$ in accordance with the degree of ionization: Let d equal the observed molecular lowering. Let n equal the number of ions resulting from the ionizing of 1 molecule of the solute. Let α equal the degree of ionization to be determined. If the ionization of 1 molecule produces n ions, we obtain in the case of this molecule $n - 1$ more particles than we had at the start. If α represents the fraction of molecules ionized, the total number of particles added by this process will be $\alpha(n - 1)$; and if we take 1 to represent the number of molecules before any ionization takes place, then the total number of particles present will be $1 + \alpha(n - 1)$. These data may be expressed thus:

$$d : 1.86 :: 1 + \alpha(n - 1) : 1.$$

From this we obtain

$$\alpha = \frac{d - 1.86}{1.86(n - 1)}.$$

Another very important method of determining the degree of ionization is the conductivity method of Kohlrausch. Before considering this method, however, it will be necessary to present certain electrochemical factors.

7. Electrochemical Units and Laws Pertaining to Conductivity.— The conductivity of a substance is its power to carry a current. The conductivity of a conductor is the reciprocal of its resistance.

The practical unit of electrical resistance is the ohm, i.e., the resistance at a temperature of 0° C. of a column of mercury 106.3 cm. long, with a cross-section equal to 1 sq. mm. This is equal to 10^9 absolute units.

The unit of current strength is the *ampere*, equal to 10^{-1} absolute units.

The unit of electromotive force is the *volt*, equal to 10^8 absolute units.

The unit of amount of electricity is called a *coulomb*. It is the amount of electricity that passes in one second of time with a current strength of 1 ampere.

All measurements of electrical energy are based on a fundamental law discovered by Ohm in 1827. This law states that the current in any circuit depends upon two variables, the voltage and the resistance of the circuit. The current is directly proportional to the voltage and inversely proportional to the resistance. It may be expressed thus: $C = E/R$.

Both the ampere and the volt are based upon another law as fundamental as Ohm's law, discovered by Faraday in 1833, which shows the relation between the current and the amount of chemical reaction that takes place when a current passes from a metallic conductor to an electrolytic conductor. This law may be expressed as follows:

(1) The amount of any substance deposited by the current is proportional to the quantity of electricity flowing through the electrolyte.

(2) The amounts of different substances deposited by the same quantity of electricity are proportional to their chemical equivalent weights.

The chemical equivalent of any element (or group of elements) is equal to the atomic weight (or sum of the atomic weights) divided by the valency. In other words, if the same quantity of electricity is passed through solutions of silver nitrate, cuprous chloride, and gold chloride, the relative amounts of hydrogen and the metals are as follows:

Electrolyte...................	$AgNO_3$	$CuCl$	$CuCl_2$	$AuCl_3$
Electrochemical equivalent H = 1;	$\dfrac{107.88}{1}$	$\dfrac{63.57}{1}$	$\dfrac{63.57}{2}$	$\dfrac{197.2}{3}$.

It has been determined by experiment that a *coulomb* liberates 0.00001036 gram of hydrogen ions; hence a coulomb will cause the separation of 0.00001036 × *a* grams of any other element, where *a* is the equivalent weight of the element.

8. Specific, Molecular, and Equivalent Conductivities.—The resistance offered by conductors depends upon two things, their nature and their form. Now, in order to compare the resistance of different substances it is necessary to use forms that are comparable. There are two forms that are generally used: one is a cube whose edge is 1 cm. long, and the other is a cylinder 1 m. in length and 1 sq. mm. in cross-section. The resistance of the second is obviously 10,000 times that of the former. When the resistance of such forms of substances is measured in ohms, it is known as the *specific resistance*. The *specific conductivity* is the reciprocal of the specific resistance.

In the case of solutions we have to choose some standard against which the solution can be measured. The conductivities of solutions are much less than those of metals; consequently we take as a standard the conductivity of a cube of the solution whose edge is 1 cm. In this case, the conductance does not depend on the whole of the material between the electrodes, but only on the ions formed by the solute. To compare different substances with respect to the conductitivy which they exhibit in solution, we must compare solutions of equivalent concentration, or solutions which contain gram-molecular weights of the substances that are capable of yielding ions bearing the same total charge of electricity. For example, H_2SO_4 ionizes into $2H^+$ and $SO^=$; therefore, 1 gram-molecule of sulphuric acid can yield ions (in case it is entirely dissociated) bearing a charge of 4 faradays. HCl, on the other hand, ionizes into H^+ and Cl^-; therefore, 1 gram-molecule of hydrochloric acid can yield ions bearing a total charge of 2 faradays. One gram-molecule of HCl, therefore, is equivalent to ½ gram-molecule of H_2SO_4. In this manner we obtain what is known as the *equivalent conductivity* of a solution. By equivalent conductivity, then, is meant the conductance, in reciprocal ohms, of a solution containing 1 gram equivalent of solute when placed between electrodes which are 1 cm. apart. This factor is usually represented by the Greek letter λ.

The *molecular conductivity* is the conductance, in reciprocal ohms, of a solution containing 1 gram-molecule (mol) of solute when placed between electrodes which are exactly 1 cm. apart. This factor is represented by the Greek letter μ.

If we let L represent the specific conductivity of a solution, and V_m the volume in cubic centimeters which contains 1 mol of solute, then

$$\mu = LV_m,$$

also

$$\lambda = LV_e,$$

where V_e is the volume of solution, in cubic centimeters, which contains 1 gram-equivalent of solute.

If C represents the concentration of a solution of an electrolyte in gram-equivalents per liter, then the last-mentioned equation becomes

$$\lambda = \frac{1000L}{C}.$$

9. Measuring the Conductivity of Solutions.—The determination of the conductivity of a solution resolves itself into the determination of the resistance, and this is effected by a modification of the well-known Wheatstone-bridge method. In the earlier methods for measuring the conducting power of solutions, a continuous current was used, but it soon developed that when such a current is passed through a solution, the electrolytes become quickly polarized, thereby increasing the resistance of the solution, and seriously affecting the result obtained. Kohlrausch[13] overcame this difficulty by using an alternating current. This type of a current from a small induction coil is sent through the solution, which is placed in a glass vessel containing two platinum electrodes. The resistance of the solution is balanced against a rheostat on a Wheatstone bridge, the point of equilibrium being determined by means of a telephone. (A galvanometer cannot be used very satisfactorily with an alternating current.)

The arrangement of the apparatus used in the Kohlrausch method is represented diagrammatically in Fig. 17, where B is a resistance box, ab is the bridge wire, and C is a cell containing the solution whose resistance is to be tested. The points d and c are connected to the small induction coil, I, which gives an alternating current. In order to determine the position of the point of equilibrium, a telephone, T, is inserted in the system. A is a small accumulator, which serves to operate the

induction coil. The point of equilibrium is ascertained by moving a slider, which is connected with the telephone, along the bridge wire until the sound of the coil is no longer audible in the telephone. According to the principle of the Wheatstone bridge,

$$\frac{C}{B} = \frac{bc}{ac}.$$

And, since the resistance, B, and the lengths, bc and ac, are easily ascertained, the resistance C can be calculated.
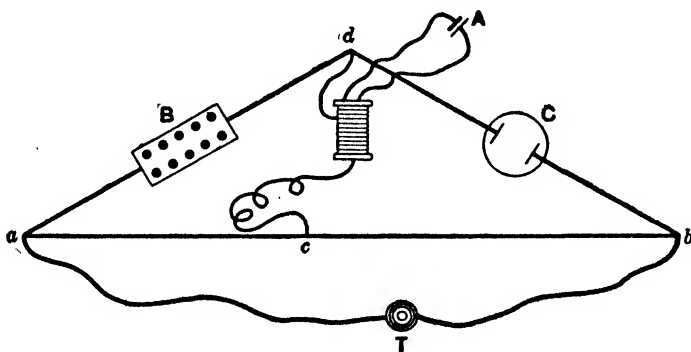

FIG. 17.

Various types of conductivity cells are employed, according as the resistance of the solution is high or low. One of the most generally serviceable types is shown in Fig. 18. The electrodes are platinum plates sealed into glass tubes, and electrical connection is made by means of mercury. These two tubes pass through a tight-fitting vulcanite cover in such a manner that the electrodes may be removed, washed, and dried, without changing their relative positions. Before use, the electrodes are coated, electrolytically, with platinum black, in order that the sharpness of the sound minimum may be intensified. It is not necessary to know the area of the electrodes, nor the distance between them, since it is possible to determine a factor by means of which the result obtained with the cell can be transformed into reciprocal ohms. This factor, which depends on the size and shape of the electrodes, and on the distance between them, is known as the *resistance capacity* of the cell, or the *cell constant*. To obtain this factor, standard solutions of known concentration must be used. The solution most commonly employed is a 0.02 molar solution of potassium

chloride. In the preparation of the standard solution, as in all other conductivity experiments, it is necessary to use especially pure distilled water, so-called *conductivity water*, since the conductivity may be appreciably altered by slight traces of impurities. The conductance of an electrolytic solution is also greatly affected by temperature; hence, it is necessary to carry out the tests at a constant temperature. This is done by placing the cell in a thermostat. Kohlrausch found the specific conductivity, L, for 0.02 KCl at 18° C., to be 0.002397, and the equivalent conductivity, $\lambda_{18°}$, to be 119.85 (reciprocal ohms).

10. Influence of Dilution on Specific, Molecular, and Equivalent Conductivity.—It has already been made clear that the conductivity of a solution does not depend on the solvent but on the solute, and that, owing to this fact, it is much more convenient to deal with solutions containing quantities of solute proportional to the respective molecular weights. The *molecular conductivity*, μ, is usually used in this connection; it is the conductivity, in reciprocal ohms, of a solution containing 1 mol of the solute when placed between electrodes exactly 1 cm. apart. We have also seen that it is equal to the specific conductivity of a solution multiplied by the volume, in cubic centimeters, that contains 1 mol of the solute; i.e., $\mu = LV_m$, where $L =$ the specific conductivity and $V_m =$ the volume in cubic centimeters that contains 1 mol of solute. As an example, the following values for the specific conductivity and the molecular conductivity of solutions of sodium chloride at 18° C., as given by Kohlrausch, may be quoted.
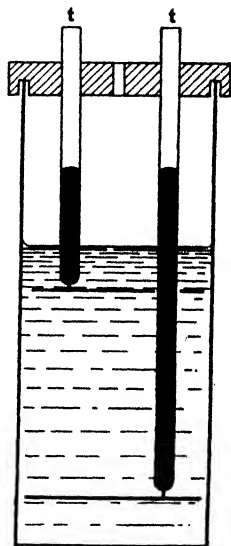
FIG. 18.

Concentration of Solution	Specific Conductivity, L	Molecular Conductivity, LV_m
1.0 molar ($V_m = 1,000$).......	0.0744	74.4
0.1 molar ($V_m = 10,000$)......	0.00925	92.5
0.01 molar ($V_m = 100,000$).....	0.001028	102.8
0.001 molar ($V_m = 1,000,000$)...	0.0001078	107.8
0.0001 molar ($V_m = 10,000,000$)..	0.00001097	109.7

Note that the molecular conductivity increases at first with dilution, and then, after a certain point, remains practically constant.

That the specific conductivity should diminish with increasing dilution is only to be expected, for its value is in each case based on the resistance between two opposite faces of a 1-cm. cube filled with the solution. Now, as the solution is gradually diluted, there will be less of the salt in this cube, and less, therefore, of the substance that acts as the carrier of the current, for the water is a non-conductor, or, at all events, a very poor conductor. It is quite natural, then, that the specific conductivity, based as it is on the consideration of a definite volume of the solution, should diminish with the dilution.

The equivalent conductivity of all electrolytes, on the other hand, increases with dilution. The extent of the effect of dilution is shown very·clearly in Table X, which gives the equivalent conductances of several typical electrolytes. The results are expressed in reciprocal ohms.

TABLE VIII

SALT SOLUTIONS

Equivalent Conductivity at 18°

Gram-equivalent per liter	KCl	CH$_3$COONa	HCl	NaOH	CH$_3$COOH
1.0	98.3	41.2	301	160	1.32
0.5	102 4	49.4	327	172	2.01
0.1	112.0	61 1	· 351	183	4.60
0 05	115.0	64 2	360	190	6.48
0 01	122 4	70.2	370	200	14.30
0 005	124 4	72 4	373	203	20.00
0 002	126.3	74 3	376	206	30.20
0.001	127.3	75.2	377	208	41.00
0 0005	128.8	75.8	....	....	57.00
0.0002	128.8	76.4	....	....	80.00
0.0001	129.1	76.8	....	....	107.00

The cases quoted in the table just given are merely instances of the behavior of aqueous solutions of electrolytes generally, and there is therefore no doubt that the equivalent conductivity increases with dilution. The extent of this increase depends on the nature of the solute, as is evident from the tabulated values given in the table.

Such substances as KCl, CH_3COONa, HCl, and NaOH show that the λ increases only very slightly in the most dilute solutions, that, in fact, it tends towards a maximum value which can be found by extrapolating to zero concentration. This may be done on the basis of an empirical rule discovered by Kohlrausch. He showed that for most electrolytes in dilute solution there is a linear relationship between the equivalent conductivity and the cube root of the concentration. This value by extrapolation is known as the *equivalent conductivity of infinite dilution*, and is indicated by the symbol λ_∞. The extrapolation principle is not applicable in cases like that of acetic acid, where λ is increasing rapidly at great dilution. In such cases other methods have been adopted.

11. Conductivity as a Means of Measuring the Degree of Dissociation.—The principal use of the conductivity method is to measure the degree of dissociation of solutions of electrolytes. Theoretically, the application of the conductivity method to the determination of the degree of dissociation is very simple. If there is no dissociation, there is no conductivity. If the dissociation is complete, the conductivity is greater than at any other dilution, since we are dealing with gram-molecular quantities. If the dilution is only partly dissociated, the conductivity will lie somewhere between zero and the maximum value. It follows, then, that the conductivity is proportional to the dissociation; therefore, the dissociation at any dilution is simply the ratio between the conductivity at the dilution in question and the conductivity at complete dissociation. If the conductivity at any dilution, V, is represented by λ_V, the conductivity at complete dissociation by λ_∞, and the percentage of dissociation by a, we have the following relation:

$$a = \frac{\lambda_V}{\lambda_\infty}.$$

The value of λ_∞ for strong electrolytes can be determined by the extrapolation method of Kohlrausch, which was given above. For KCl, for instance, at 18° C., it has been calculated to be 130.10. From the table on page 168 we see that a solution of KCl containing 0.01 gram-equivalents per liter has the equivalent conductivity of 122.4 reciprocal ohms. In this solution, therefore, the degree of dissociation is

$$a = \frac{122.4}{130.1} = 94.08 \text{ per cent.}$$

The values of λ_∞ at 18° for potassium chloride, sodium hydroxide, hydrochloric acid, acetic acid, and sodium acetate are 130.1, 217.5, 383.3, 351.7, and 77.2, respectively. On the basis of these numbers, and of the figures given in Table X, the following values have been calculated for a few selected concentrations:

<div align="center">

TABLE XI

DISSOCIATION VALUES AS DETERMINED BY CONDUCTIVITY

</div>

Gram-equivalent per liter	KCl	CH₃COONa	HCl	NaOH	CH₃COOH
1 00	0.76	0.53	0.79	0 73	0.004
0.50	0 79	0.64	0 85	0 79	0.006
0.10	0.86	0 79	0.95	0.84	0.013
0.01	0.94	0 91	0.96	0.92	0.041
0.001	0.98	0 97	0.98	0.96	0.117

Table XI shows very plainly that potassium chloride, sodium acetate, hydrochloric acid, and sodium hydroxide are highly ionized in dilute solution. From data not given in the table a similar conclusion may be drawn for nitric acid, sulphuric acid, hydrobromic acid, hydriodic acid, and potassium hydroxide, and for all sodium and potassium salts of monobasic acids. Acetic acid, on the other hand, is only slightly ionized in dilute solutions, and in this respect is typical of many monobasic organic acids. Ammonia is likewise only slightly ionized.

12. Relative Speed of Ions.—Another important property of electrolyties is the relative speed with which the ions move in solution. Faraday's law states that when an electric current is passed through a solution of electrolytes, the ions move through the solvent towards the poles, the cations to the cathode and the anions to the anode. It has long been known, however, that the ions do not all move at the same rate of speed; in fact, Hittorf studied this phenomenon before the modern theory of ionization was developed (1853–59). Later Kohlrausch determined the relative rate of speed at a definite temperature, 18° C., and obtained the figures given in Table XII for a constant gradient.

The rate of speed was found to depend on the atomic weight of the ions, the degree of hydration, and the influence of other ions. (For

experimental methods of determining the relative speed of ions, consult larger texts, e.g., Jones' "Elements of Physical Chemistry," 4th ed., p. 379. (The Macmillan Co.)

TABLE XII

Cations +			Anions −		
Ion	Atomic Weight	Relative Speed	Ion	Atomic Weight	Relative Speed
H	1	318	OH	17	174
K	39	64.6	½SO₄	48	68
NH₄	18	64 4	Br	80	67
½Ba	68.5	55	I	127	66 5
½Sr	43.7	51	Cl	35.5	65.5
½Ca	20	51	½C₂O₄	44	63
½Mg	12	45	CH₃COO	59	33.7
Na	23	43.5			
Li	7	33.4			

13. Hydration of the Ions and the Velocities with which They Move.—Certain apparent discrepancies presented themselves in the velocities of the different ions, which for a time, could not be explained. For example, the lithium ion, under the same driving force, moves more slowly than the potassium ion, yet it has a smaller volume and a smaller mass. It was not until it was shown that the lithium ion is more strongly hydrated than the potassium ion that this fact could be explained. A relation between the migration velocities of the ions and their hydrating power was worked out by Jones and Pierce.[14] A part of their discussion is repeated here to bring out the point in question:

"The velocities of the ions in moving through any given medium may be assumed to vary inversely as their mass, the driving force being constant. Their velocities would also vary inversely as their volumes. Mass being constant, we should expect the ions with the smallest atomic volumes to move the swiftest under a constant driving force, while the facts are often the opposite. Leaving out of account the hydrogen and hydroxyl ions, potassium, rubidium, and caesium have very great velocities and the largest volumes; while the ions of iron and copper have the smallest volumes and very small velocities. The mean-

[14] Carnegie Inst. of Wash., Pub. No. 180, pp. 84–86.

ing of this apparent discrepancy can be seen at once by comparing the atomic volume curve and the migration velocity curve.

"The elements with the smallest atomic volumes have the greatest hydrating power. When these atoms appear as ions they frequently have the smallest velocities. Therefore, the ions with smallest velocities have the greatest hydrating power. To discuss the relation somewhat in detail, the atomic volumes of potassium, rubidium, and caesium increase rapidly with increasing atomic weight, and their salts generally crystallize without water. The atomic volume of sodium and lithium are less than half that of potassium, and yet their ionic velocities are only about two-thirds that of potassium. It will be recalled that salts of sodium and lithium may crystallize with 2 or 3 molecules of water. We may therefore assume that the increase in the volume and mass of the lithium and sodium ions, due to the formation of a hydrate, decreases the velocity of these ions below that of potassium.

"The small velocity of the lithium was, as we have seen, for a long time unexplained. Lithium has a volume only about half that of sodium, and the largest ascertained amount of water with which the salts of lithium crystallize is 3. The maximum amount for many of the salts of sodium is 2. The lithium ion is, in general, more hydrated than the sodium ion, and its velocity is therefore decreased by more hydration. Notwithstanding its smaller volume and lighter mass, on account of its greater hydration lithium moves with about the same velocity as sodium.

"The calcium atom is slightly larger than sodium, but the calcium ion has considerably smaller velocity. This is undoubtedly due primarily to its much greater hydrating power. Within this group the atomic volumes increase with increasing atomic weight. The velocities of calcium and strontium, with many salts crystallizing with 6 molecules of water, are approximately equal to that of barium. Many of the salts of barium crystallize with 2 molecules of water or water-free. The larger mass of the barium atom itself diminishes the velocity. Magnesium, with about half the volume of calcium, has nearly the same velocity, due to its greater hydrating power. Cobalt, nickel, and copper atoms have nearly the same volumes and approximately the same hydrating power. Their ions have approximately the same velocities.

"The atomic volumes of chlorine, bromine, and iodine are approximately the same. If, as ions, they hydrate at all we should expect the

same order of hydration for all three, as has been made probable. We should expect them to have velocities of the same order of magnitude, and such is the fact.

"The silver ion is the only well-established exception. Silver has a small volume and many of its salts crystallize without water. Although it has a small volume, it apparently has but little hydrating power. Notwithstanding its considerable mass, with its small volume and small hydrating power we should expect it to have a fairly high velocity. The fact is, the velocity of the silver ion is slightly less than that of chlorine, bromine, and iodine.

"The general truth of the relation that the ions with the smallest velocities have the greatest hydrating power is, then, established by the facts, the great hydrating power being one of the factors conditioning the small velocity."

14. Effect of Temperature on the Migration Velocity of Ions.— Increase in temperature, according to the kinetic theory and laws of energy, will increase the speed of ions, provided, of course, that dissociation is complete. Partially dissociated salts are more completely ionized by increase in temperature. The more highly hydrated the ion, the greater is its temperature coefficient; hence, for equal augmentation of temperature, different ions increase in speed according to their degree of hydration. This is explicable on the assumption that elevation of temperature favors the disruption of hydrated complexes and decreases the size of the ion, thereby reducing the frictional resistance to its passage through the solvent.

15. The Ionization of Water.— The solutions that we have been considering in the foregoing discussions have all been aqueous. This is natural since water is by far the most important solvent. (See Chapter VI of this text.)

In our consideration of conductivity we have regarded the conductivity of the solution as depending entirely on the ions of the dissolved electrolyte. It has been found, however, that pure water is not quite a non-electrolyte. A small quantity of H^+ ions and OH^- ions is always present in the purest water that it is possible to prepare. The degree of ionization is largely dependent on the temperature. At $0°$ C. the specific conductance of pure distilled water is 0.01×10^{-6}, and at $25°$, 0.17×10^{-6} reciprocal ohms. From the value of the specific conductance, the degree of ionization of pure distilled water can be calculated. The conductance of a cube of water of 1 cm. sides, at

18°, is 0.043 $\times$ 10^{-6} reciprocal ohms. The conductance of 1000 c.c., when placed between the electrodes 1 cm. apart, would be 1000 times this figure, or 0.043 $\times$ 10^{-3}. If 1 liter of water contained 1 gram-molecule (17 grams) of H$^+$ and OH$^-$ ions, the equivalent conductivity would be 492, that is, it would be equal to the sum of the ionic conductivities (318 and 174). The value found, however, is only 0.043 $\times$ 10^{-3}; hence the water contains only

$$\frac{0.43 \times 10^{-3}}{492} = 0.87 \times 10^{-7} \text{ gram-equivalents per liter.}$$

In other words, at this temperature, 1 gram-equivalent of H$^+$ and OH$^-$ is contained in about 11$\frac{1}{2}$ million liters of water.

Another noteworthy fact in connection with the ionization of water is that a liter of pure water at about 22° C. will contain 10^{-7} moles of H$^+$ ions and 10^{-7} of OH$^-$ ions; in other words, the concentration of each of the ions is 10^{-7}. The most surprising thing, is, however, that the concentration of the two ions, H$^+$ and OH$^-$, is always a constant. In pure water, or, in fact, in any neutral solution, the product is 10^{-7} $\times$ 10^{-7}, or 10^{-14} at 22°. In an acid solution the concentration of H$^+$ will, of course, be larger, but the corresponding concentration of OH$^-$ will be smaller; for instance, if the H$^+$ is 10^{-3}, then OH$^-$ will be 10^{-11}, so that the product is still 10^{-14}. In basic solutions we have the opposite condition, i.e., the OH$^-$-ion concentration is large and the H$^+$-ion content correspondingly small; e.g., if OH$^-$ is 10^{-4}, then the H$^+$ will be 10^{-10}, so that here again the product of the two ions is 10^{-14}. It is evident, then, that if we say a solution has an H$^+$ content of 10^{-9}, this is equivalent to saying that the OH$^-$ content is equal to 10^{-5}, since the two are always equal to 10^{-14}. The explanation of this peculiar phenomenon will be considered after the Law of Mass Action has been discussed. (See Chapter IX.)

16. Amphoteric Electrolytes.—It has already been stated that acids are compounds which in the presence of a dissociating solvent yield H$^+$ ions; and that bases are compounds which in the presence of a dissociating solvent yield OH$^-$ ions. It might be inferred from these statements that a compound which under one set of conditions yields H$^+$ ions would invariably yield H$^+$ ions on dissociation; and that, similarly, a compound which under one set of conditions would

split off OH⁻ ions would under all conditions yield the OH⁻ ions. Such is not the case, however. There are many compounds which under one set of conditions yield H⁺ ions, while under other conditions they are known to dissociate OH⁻ ions. Such compounds are known as amphoteric electrolytes, or ampholytes. There is a fairly large number of these substances, both in general chemistry and among the carbon compounds.

Some amphoteric electrolytes yield H⁺ ions and OH⁻ ions at the same time. Water itself may be regarded as a typical representative of this class, since the fraction that ionizes must yield both ions:

$$H_2O \rightleftarrows H^+ + OH^-.$$
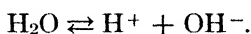
We have seen that water has a very low coefficient of dissociability; this is true of all amphoteric electrolytes. In the case of water, it is due to the fact that liquid water consists mostly of dihydrol, H_4O_2, with some trihydrol, H_6O_3, and a lesser amount of hydrol, H_2O. All of the ionization, of course, comes from the hydrol. If this fact is accepted, the reason for the increase in ionization with increase in temperature becomes apparent, for as the temperature increases the amount of hydrol increases, and consequently the number of ions.

Of the many examples of amphoteric substances in the field of general chemistry, only a few will be mentioned. We have, for instance, zinc hydroxide, which in the presence of an acid yields OH⁻ ions as follows:

$$Zn(OH)_2 = Zn^{++}, OH^-, OH^-;$$
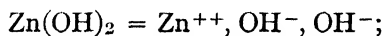
while in the presence of an alkali, it dissociates thus:

$$Zn(OH)_2 = ZnO^=, H^+, H^+.$$
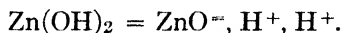
In the first case zinc hydroxide dissociates as a base; in the second, as an acid. For this reason zinc hydroxide dissolves in an excess of sodium hydroxide almost as readily as in hydrochloric acid.

Aluminium hydroxide, $Al(OH)_3$ is another example. Its amphoteric nature is shown by the fact that it dissolves in acids to form aluminium salts, e.g., $AlCl_3$; or in alkalies to form aluminates, e.g., $NaAlO_2$. This dual reaction of aluminium hydroxide can be seen from the following:

$$3H^+ + AlO_3^= \rightleftarrows Al(OH)_3 \rightleftarrows Al^{+++} + 3OH^-.$$

Among the most important amphoteric electrolytes, from a biochemical standpoint, are the amino acids, the cleavage products of the proteins. They contain both carboxyl (COOH) and amino (NH_2) groups, and can therefore dissociate either as acids or as bases. Glycine, amino acetic acid ($H_2NCH_2 \cdot COOH$) will serve to illustrate the point. It can form both a sodium salt, $H_2N \cdot CH_2 \cdot COONa$, and a hydrochloride, $HCl \cdot H_2N \cdot CH_2 \cdot COOH$. In strong acids it behaves as a base; in strong alkalies, as an acid. In neutral solutions it is a feeble electrolyte, and is partially dissociated into H^+ and an anion:

$$H_2N \cdot CH_2 \cdot COOH \rightleftarrows H_2N \cdot CH_2 \cdot COO^- + H^+.$$
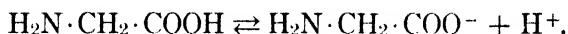
It also dissociates partially into OH^- and a cation:

$$HO \cdot H_3N \cdot CH_2COOH \rightleftarrows HO^- + {}^+H_3N \cdot CH_2 \cdot COOH.$$

The amino acids in conjunction with certain other amphoteric electrolytes play a very important part in maintaining the normal reaction of the tissues.

CHAPTER IX

CHEMICAL EQUILIBRIUM AND THE LAW OF MASS ACTION

1. Reversible and Non-reversible Reactions.—When substances take part in a reaction, some of the molecules are decomposed while others are being formed. This process may continue until one or more of the reacting compounds has entirely disappeared from the field of action, in which case the reaction is said to be *complete*. This type of reaction is also often described as *non-reversible*. For example, if one mixes solutions containing equimolecular amounts of silver nitrate and sodium chloride, reaction will take place, and will continue until all of the silver has combined with chlorine. The silver chloride, being insoluble, will precipitate, and hence will be removed from the other reacting particles.

Another slightly different example is that of aqueous solutions of hydrochloric acid and sodium hydroxide. In this case the reaction will continue until either one or the other disappears; they cannot exist together in the same solution. These changes are represented by the following equations:

$$AgNO_3 + NaCl = AgCl + NaNO_3,$$

and

$$HCl + NaOH = NaCl + HOH.$$
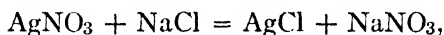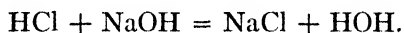
It is evident that such complete reactions are *non-reversible*.

There are, however, many reactions which do not take place completely in one direction, as represented in the foregoing equations. They do not proceed until one or the other of the substances has disappeared, but instead the reaction continues until a condition is reached in which the changes in one direction are exactly balanced by those in the other. At this point, a condition of equilibrium is said to have been established between the reacting substances.

2. The Law of Mass Action as Applied to Chemical Equilibrium.— Suppose that we bring together in any homogeneous gas or liquid system two substances, A and B, which are capable of entering into

chemical reaction with each other to form two other substances, C and D, as indicated by the equation:

$$A + B = C + D.$$
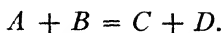
As the reaction proceeds, the concentrations of A and B will tend to decrease, and those of C and D to increase. At first the reaction will proceed rapidly from left to right, and then more slowly, until finally a condition will be attained in which no change in the molecular concentrations of the reacting substances is observable, i.e., from all appearances the reaction is at a standstill. In like manner, if we were to start with the two substances C and D at similar equivalent concentrations as were used for A and B, we should observe that the reaction would proceed with gradually diminishing speed from *right to left*. Again, the reaction would continue until a state was attained in which no concentration change was observable. It must not be assumed that the reaction is a one-way process depending on which couple we start with: as soon as A and B start to react to produce C and D, C and D also start to react to produce A and B and *vice versa*. Furthermore, the concentrations of the four substances under consideration, at the point of equilibrium, have been found to be independent of the direction from which this condition is approached. In other words, we cannot change the concentration of any substance in an equilibrium mixture by changing the direction of approach to equilibrium. That is, whether we start with a moles of A and b moles of B or c moles of C and d moles of D, the concentrations of all four substances at the stage when no further concentration change is observable will have the same respective values, no matter how the reaction started. It must not be assumed, however, that when the reaction has apparently ceased, i.e., when a condition of equilibrium is attained, all reaction has ceased, but rather that the velocity of change in one direction has become equal to that of the opposite direction, with the result that no further total change can be observed, i.e., it is only the total (apparent) change which becomes zero.

The concentrations of the reactants when the condition of equilibrium is attained are called their *equilibrium concentrations*. It is customary to express the concentrations of the reacting substances by the use of brackets; thus, the rate of the reaction:

$$A + B \rightleftarrows C + D$$
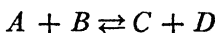
in which A and B react to produce C and D, while C and D react to form A and B, is expressed as:

$$K(A) \cdot (B) \rightleftarrows K'(C) \cdot (D),$$

or

$$K(C)_A \cdot (C)_B \rightleftarrows K'(C)_C \cdot (C)_D.$$

Where K and K' are the *equilibrium constants* of the two reactions, respectively. It should also be noted that the ratio of these two quantities defines the composition of the system in equilibrium; if one reaction proceeds twice as fast as the other, then, in order to bring up the slower reaction to that of the faster, as must be the case in equilibrium, the concentration of the reacting substances in its case must be correspondingly increased. It is evident, then, that in any chemical equilibrium in a perfect gas or a dilute solution at a given temperature, the equilibrium concentrations, $(C)_A$, $(C)_B$, $(C)_C$, $(C)_D$, are always so related that they fulfill the following condition:

$$\frac{(C)_A \times (C)_B}{(C)_C \times (C)_D} = K,$$

indicating that no matter how the concentrations A, B, C, D are varied, reaction will take place in one direction or the other until the concentrations have become adjusted so that the "equilibrium constant," K, remains unchanged. This relation is the *law of mass action* as applied to chemical equilibrium.

Let us consider, as an example of the application of the law of mass action, the reaction that occurs between alcohols and organic acids. These substances react to form esters and water, but esters and water react to form alcohols and acids, and in neither case does the reaction take place completely; i.e., the acids and the alcohols do not entirely disappear, but only a fraction of them, no matter how long they are left in contact. This is due to the fact that the ester and water formed react to form the original acid and alcohol. In short, the reaction tends to be reversed. It is evident, then, that the equilibrium equations given above are applicable to these substances.

In the case of ethyl alcohol and acetic acid, which react to form ethyl acetate and water, and *vice versa*, the equilibrium equation is

$$\underset{\text{ethyl alc.}}{C_2H_5OH} + \underset{\text{acetic acid}}{CH_3COOH} \rightleftarrows \underset{\text{ethyl acetate}}{CH_3COOC_2H_5} + H_2O$$
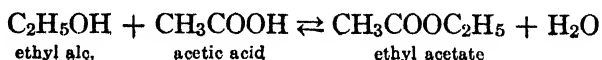
or, expressed according to the law of mass action,

$$\frac{(C)_{acid} \times (C)_{alcohol}}{(C)_{ester} \times (C)_{water}} = K.$$

The value of the equilibrium constant, K, must be determined by experiment; in the case just mentioned it is equal to $\frac{1}{4}$. If equimolecular proportions of acetic acid and ethyl alcohol are mixed together, one-third of each of them remains unchanged, no matter how long they are left in contact. It follows, therefore, that if the initial concentration of acid and alcohol is 1, we shall have at equilibrium,

$$K = \frac{\frac{1}{3} \times \frac{1}{3}}{\frac{2}{3} \times \frac{2}{3}} = \frac{1}{4}.$$

Now, it does not matter in what proportion we mix the four reactants, the same condition of equilibrium is always attained. In other words, the product of concentrations of acid and alcohol divided by the products of concentrations of the ester and water will be equal to $\frac{1}{4}$. Suppose, for instance, that when equilibrium is attained, a further quantity of acetic acid is added. We increase one of the components of the mixture in so doing; the product of concentrations in the numerator is then increased, and, in order to maintain the above equation, the value of the denominator must also be increased. That is, there must be a further combination of alcohol and acid to form ester and water. By increasing the concentration of acid sufficiently, practically all the alcohol can be converted into ester. This brings out the principle that *incomplete reactions may be made to take place more completely by adding an excess of any one of the reactants.*

3. Importance of the Law of Mass Action in Physiological Processes.—The application of the law of mass action in the explanations of biochemical processes is very extensive. Most of the chemical reactions that occur in the process of digestion and in living cells are reversible. Moreover, the majority of these reactions take place very slowly unless the process is augmented by the assistance of a *catalyst*. A catalyst is a substance that changes the velocity of a chemical reaction which is already taking place. The basis of the study of reaction velocity is furnished by the law of mass action, which we have employed to investigate the conditions of equilibrium. The *biochemical catalysts*, or *enzymes*, as they are called, are manufactured in the living

cells. These enzymes accelerate the velocity reaction of chemical processes existing in the organism, in a similar manner to the action of inorganic catalysts in general chemical reactions. For example, an enzyme present in the pancreatic juice, called *lipase*, accelerates the hydrolysis of such esters as neutral fats, which are triatomic esters of glycerol and fatty acids, such as stearic ($C_{17}H_{35} \cdot COOH$), palmitic ($C_{15}H_{31} \cdot COOH$), and oleic ($C_{17}H_{33} \cdot COOH$).

$$C_3H_5(OOC \cdot C_{17}H_{35})_3 + 3H_2O \overset{+\text{Enzyme}}{\rightleftarrows} 3(C_{17}H_{35}COOH) + C_3H_5(OH)_3$$
<div style="text-align:center">Stearin (ester)   Stearic acid   Glycerol</div>
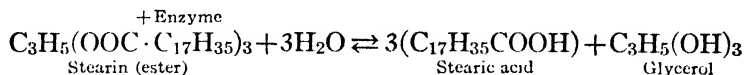
As will be made clearer in Chapter XII, the amount of water present is the principal determining factor which controls the direction in which the reaction (in this case) proceeds to the greatest extent. Under the conditions which develop during digestion, the reaction proceeds from left to right until nearly all the neutral fat has been broken down into glycerol and fatty acids, because of the preponderance of water; but if we start with a mixture of fatty acids, glycerol, enzyme, and just enough water to permit the enzyme to act, the reaction will proceed from right to left, i.e., neutral fat will be synthesized. This is known as the reversible action of enzymes. Such reactions are continually taking place in metabolism, and therefore are very important. Just how the small percentage of water is brought about in certain places in the organism, thereby permitting synthesis to take place in living tissues, will be more easily understood after we have considered the properties of the colloidal state of matter (Chapter XI), velocity of reaction (Chapter XII), and permeability of cell membranes (Chapter XIII).

Another topic that we shall consider later is "the maintenance of neutrality in organisms," i.e., the maintenance of a practically constant balance between the H^+ and OH^- ions. The explanation of this topic is likewise fundamentally dependent on the law of mass action.

In short, it is impossible to understand most biochemical processes without a knowledge of the principle of the law of mass action.

4. Equilibrium between Ionized and Un-ionized Solute: Ostwald's Dilution Law.—We have already seen that the law of mass action is applicable to chemical equilibria, both in perfect gases and in dilute solutions. We shall now see that it applies equally to electrolytes in equilibria. Acids, bases, and salts, when dissolved in water,

are ionized; the ions, at the same time, recombine to form neutral molecules. It is evident, then, that equilibrium will ensue when the speeds of the two opposing reactions—ionization and molecule formation—are equal. Consider the ionization of acetic acid, which when dissolved in water dissociates according to the equation

$$CH_3COOH \rightleftarrows CH_3COO^- + H^+.$$

If we apply to this equilibrium the law of mass action, the condition for equilibrium is

$$\frac{(CH_3 \cdot COO^-) \times (H^+)}{(CH_3 \cdot COOH)} = K.$$

Suppose that 1 gram-molecule of acetic acid is dissolved in water, that the solution is then diluted to V liters, and that α represents the degree of dissociation; the amount ionized at equilibrium will be $(1 - \alpha)$ and the amount of each of the ions will be α gram-equivalents. The concentration of un-ionized molecules will therefore be $\dfrac{(1 - \alpha)}{V}$, and that the concentration of each of the ions $\dfrac{\alpha}{V}$. Applying the law of mass action, we have

$$\frac{(\alpha)^2}{V} = K\frac{(1 - \alpha)}{V},$$

and from this,

$$K = \frac{\alpha^2}{(1 - \alpha)V},$$

where K is the equilibrium constant.

This result was worked out by Ostwald,[1] and is known as the *Ostwald dilution law*. It is found experimentally to apply only to weak acids and bases. A different law, which is not dependent on the law of mass action, applies to salts, strong acids, and strong bases.

Ostwald's dilution law has been tested with numerous acids, bases, and salts. The results of three typical cases, acetic acid, ammonium, and silver nitrate, are given:

[1] Zeit. phys. Chem., **2**, 36, 1888.

TABLE XIII

V	K (Acetic Acid)	$K(NH_3)$	$K(AgNO_3)$
8	0.0000180	0.000023	
16	0.0000179	0.000023	0.253
32	0.0000182	0.000023	0.191
64	0.0000179	0.000023	0.127
128	0.0000179	0.000023	0.122
256	0.0000180	0.000023	0.124
512	0.0000180	0.000024	0.125
Mean	1.80×10^{-5}	2.3×10^{-5}	

From Table XIII, we see that the values for the salt, silver nitrate, are not constant, but that the values for such weak acids and bases as acetic acid and ammonia are so constant that we are justified in making the following generalization: In those cases in which the dilution law applies, the *dissociation constant*, K, is independent of dilution, and can be used in comparing the strength of the electrolyte concerned.

It is instructive to compare the values of K obtained for different organic acids. A few typical cases obtained at 25° C. are given below:

TABLE XIV

Acid	K
Acetic Acid..........................	0.000018
Monochlor-acetic acid..................	0.001550
Dichlor-acetic acid....................	0.051400
Trichlor-acetic acid...................	1.210000
Benzoic acid..........................	0.000060
Salicylic acid........................	0.001020
Formic Acid..........................	0.000214

Note how the introduction of chlorine into the acetic acid molecule increases the value of the dissociation constant; it is also interesting to note the effect of a hydroxyl radical in the benzene ring. Compare the value of K for salicylic acid with that for benzoic acid.

5. Equilibrium Relations of Water.—Water, as we have learned (p. 173), is ionized to a slight degree into H^+ ions and OH^- ions. If we apply the law of mass action to the equilibrium

$$H_2O \rightleftharpoons H^+ + OH^-,$$
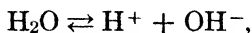
we obtain

$$\frac{C_{H^+} \times C_{OH^-}}{C_{H_2O}} = K,$$

where C_{H^+}, C_{OH^-}, and C_{H_2O} are the concentrations of the H^+ ions, the OH^- ions, and non-dissociated water, respectively. The concentration of H^+ and OH^- is extremely small; it is only 1×10^{-7} mole per liter at 20° C. The concentration of the un-ionized fraction is 1000/18, or 55 molar. This may also be written 0.55×10^2. Substituting values in the above equilibrium, we obtain for K

$$\frac{10^{-7} \times 10^{-7}}{0.55 \times 10^2} = 2 \times 10^{-16}.$$

Now, as the quantity of dissociated water is so very small compared with the total quantity of water, it is clearly justifiable to put (H_2O) = 1 in the mass action equation, which then becomes

$$C_H \times C_{OH} \quad \text{or} \quad (H^+) \times (OH^-) = K.$$

Stated in words, the product of concentration of hydrogen ions and hydroxyl ions is always equal to a constant quantity, K_w. The value of K_w at 22° C. (average room temperature in this country) is 1×10^{-14}.

At other temperatures the value of K_w will be quite different, since this constant depends only on temperature. Table XV will bring out this point quite clearly:

<div align="center">

TABLE XV

</div>

$0° = 0.10 \times 10^{-14}$	$50° \times \ 4.5 \times 10^{-14}$
$18° = 0.86 \times 10^{-14}$	$75° = \ 16.9 \times 10^{-14}$
$22° = 1.00 \times 10^{-14}$	$100° = \ 48.0 \times 10^{-14}$
$25° = 1.20 \times 10^{-14}$	$218° = 461.0 \times 10^{-14}$

6. Strength of Acids and Bases.—There are several methods by which the relative strengths of acids and bases can be estimated. One of the methods which has proved to be of great value, especially for weak acids and bases, is based on the relative dissociation constant.

As already stated, $K = \dfrac{\alpha^2}{(1 - \alpha)V}$, so that if we are considering two acids, for instance, which are feebly dissociated, for which therefore $1 - \alpha$ is practically 1, and if we compare the two acids at the same concentration and temperature, we have

$$K_1 = \frac{\alpha_1{}^2}{V} \quad \text{and} \quad K_2 = \frac{\alpha_2{}^2}{V}.$$

Now, if we remember that the value of K for an electrolyte is an expression of its inherent ability to dissociate into ions, and that any electrolyte which yields hydrogen ion (H^+) as one of the direct products of its ionization is called an *acid*, it then becomes evident that the extent to which such an electrolyte dissociates is in a sense a measure of its strength as an acid. In other words, the value of K is a measure of the strength of an acid. Basic properties similarly depend on the OH^- ion, so that the dissociation constant, K, is also a measure of the strength of the base.

We can also express the relation between the H^+-ion concentration, the degree of dissociation of an acid, and the value of K in another manner, which has been found useful in many cases.

In the case of weak acids, if we represent the un-ionized acid molecule by (HA) and the anion by (A'), we may write the mass law equation thus:

$$\frac{(A')}{(HA)} = \frac{K}{H^+},$$

and, since (HA) = (total acid) − (A'), we obtain

$$\frac{(A')}{(\text{Total acid}) - (A')} = \frac{K}{(H^+)} \quad \text{or} \quad \frac{A'}{\text{Total acid}} = \frac{K}{K + (H^+)} = \alpha.$$

This expression may be transformed into

$$\frac{1}{(H^+)} = \frac{1}{K} \cdot \frac{\alpha}{1 - \alpha},$$

and if the logarithm of each side is taken we get

$$\log \frac{1}{(H^+)} = \log \frac{1}{K} + \log \frac{\alpha}{1 - \alpha}.$$

The relative strength of acids can also be determined from their

catalytic effect on the rates of certain reactions, such as the hydrolysis of esters or the inversion of sugars; however, it would lead us too far astray if we were to discuss the details of this method.

We have already learned that Arrhenius (see Chapter VIII) devised a method for the estimation of the relative strengths of acids and alkalies from their electrical conductance. The results obtained with this method are comparable with those obtained with the others to which reference has just been made.

7. Ionization of Polybasic Acids.—The presence of considerable amounts of such polybasic acids as carbonic acid, uric acid, phosphoric acid, etc., in physiological media, and the rôle that these substances play in life processes make the consideration of their ionization important.

Polybasic acids ionize in what might be termed "progressive steps." Thus, when a dibasic acid, such as uric acid or carbonic acid, is dissolved in water, ionization takes place as follows:

$$(1) \quad H_2A \rightleftarrows H^+ + HA'.$$
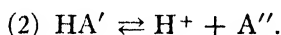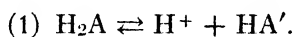
$$(2) \quad HA' \rightleftarrows H^+ + A''.$$

It should be clear that the law of mass action applies to each of these equilibria of weak dibasic acids, and so we obtain the expressions

$$(1) \quad \frac{(H^+) \times (HA')}{(H_2A)} = K_1,$$

and

$$(2) \quad \frac{(H^+) \times (A'')}{(HA')} = K_2.$$

The separation of the first hydrogen ion of a polybasic acid is called "primary" ionization; the second to separate is called "secondary"; and the third to separate is called "tertiary." This progressive ionization of polybasic acids is shown by the following scheme for the tribasic acid, phosphoric acid:

$$H_3PO_4 \rightleftarrows H^+ + H_2PO_4^- \quad \text{(primary)}$$
$$\text{⇅}$$
$$H^+ + HPO_4^- \quad \text{(secondary)}$$
$$\text{⇅}$$
$$H^+ + PO_4^{=} \quad \text{(tertiary)}.$$
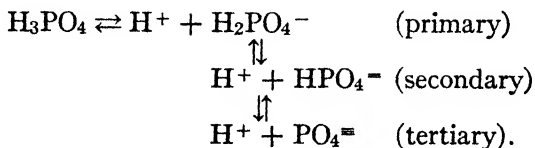
The value of K is different for each of the above stages.

The primary ionization of polybasic acids predominates in relatively concentrated solutions; the secondary, in moderately dilute solutions; while the tertiary stage is only obtained in very dilute solutions.

Table XVI gives the ionization constants of a number of polybasic acids:

TABLE XVI

Acid	K'_1	K_2	K_3
Phosphoric acid.............	$9 \ 0 \times 10^{-3}$	8.8×10^{-8}	$3 \ 6 \times 10^{-13}$
Carbonic acid..............	$3 \ 3 \times 10^{-7}$	6.0×10^{-11}	
Uric acid	$2 \ 0 \times 10^{-6}$	2.0×10^{-9}	
Tartaric acid...............	9.7×10^{-4}	4.5×10^{-5}	
Oxalic acid	3.8×10^{-2}	4.9×10^{-5}	

8. Equilibrium between Electrolytes with a Common Ion.—The reaction between the undissociated molecule of a weak acid and its ions is, as we have seen, a reversible one, and it is consequently possible to shift the equilibrium in either direction by varying the conditions. We saw in the reversible reaction between ethyl alcohol plus acetic acid and ethyl acetate plus water, that the greater the amount of alcohol the smaller was the corresponding concentration of acetic acid. We have also seen that acetic acid ionizes according to the equation

$$H \cdot C_2H_3O_2 \rightleftarrows H^+ + C_2H_3O_2^-,$$

and that the ions and the undissociated molecules are in equilibrium as represented by the equation

$$\frac{(H^+) \times (C_2H_3O_2^-)}{(H \cdot C_2H_3O_2)} = K.$$

By adding sodium acetate, which, like most salts, is highly ionized, to an acetic acid solution, we introduce into the system a large extra supply of acetate ions, $C_2H_3O_2^-$. The effect of this, by the law of mass action, is to lessen the amount of dissociated acetic acid. A 0.1M solution of acetic, by the addition of a little sodium acetate, may have its dissociation into $H^+ + C_2H_3O_2^-$ so far reduced as actually to become alkaline to such an indicator as methyl orange. Of course, its

total acidity, as measured by a strong alkali that it can neutralize in titration, is unaffected.

Perhaps we can make the effect of adding sodium acetate to acetic acid solution clearer by using figures in the above equation. Let us take molar solutions of acetic acid and sodium acetate. Molar acetic acid is 0.0042 (= 0.42 per cent) ionized. The concentration of each ion will therefore be 0.0042, and that of the un-ionized portion will be 0.9958. Substituting in the above equation, we have

$$\frac{0.0042 \times 0.0042}{0.9958} = 1.8 \times 10^{-5}.$$

Molar sodium acetate is 0.53 (53 per cent) ionized; therefore, the addition of 1 mole of this salt to the molar solution of acetic acid will increase the acetate-ion concentration by 0.53, making the total concentration of this ion 0.5342. The relation will then be

$$\frac{0.0042 \times 0.5342}{0.9958} > 1.8 \times 10^{-5}.$$

But change of concentration is not supposed to change the value of the dissociation constant for any length of time; therefore, a readjustment of the reactants will immediately take place and the original ratio will again hold. This means that a sufficient number of ions will combine to form un-ionized molecules to reëstablish the original equilibrium equation. If we let X equal the fraction of a mole thus combined, it is evident that the concentration of each ion will be decreased by X moles, and the concentration of the un-ionized part will be increased by the same amount. We shall then have

$$\frac{(0.0042 - X) \times (0.5342 - X)}{(0.9958 + X)} = 1.8 \times 10^{-5}.$$

From this, X is found to be 0.004166 mole. Now, by substituting this figure in the equation, we get

$$\frac{0.000034 \times 0.530034}{0.999966} = 1.8 \times 10^{-5}.$$

The same principle may be applied with other weak acids, and, of course, with weak bases. In general, then, the degree of ionization of

any weak acid or base is always decreased by the addition of salts having an ion in common with itself. This fact is of great importance in biology.

9. Mixture of Electrolytes with no Common Ion.—The equilibrium in a mixture of two electrolytes without a common ion can be calculated when the concentrations and dissociation constants are known, but the calculation is rather complicated. If two strong electrolytes (highly dissociated), such as NaCl and KBr, are dissolved in the same solvent, the following products will be obtained: K^+, Cl', Na^+, Br', and small amounts of KBr, NaBr, KCl, and NaCl. In other words, there will be an interaction between the ions; but as the salts are all highly dissociated, the mutual effect will be very small. At great dilution, only ions will be present.

Quite different results are obtained, however, if one of the products of the reaction is a weakly dissociated compound. Take hydrochloric acid and sodium acetate, for example; both of these electrolytes are highly dissociable, but one of the products of reaction, namely, acetic acid, is, as we know, weakly dissociated. When we mix hydrochloric acid and sodium acetate, we first get HCl, CH_3COONa, H^+, Na^+, Cl', CH_3COO', CH_3COOH, and NaCl; then from these we get the following equilibria:

$$HCl \rightleftarrows H^+ \qquad + Cl'$$

$$NaC_2H_3O_2 \rightleftarrows C_2H_3O_2{}^- + Na^+$$
$$\updownarrow \qquad\quad \updownarrow$$
$$HC_2H_3O_2 \qquad NaCl$$
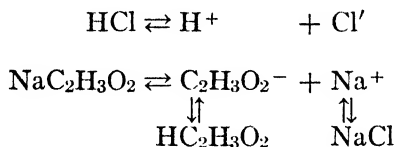
This means that the addition of sodium acetate to hydrochloric acid effects a removal of hydrogen ions through the introduction, as a product of reaction, of a weakly dissociable acid. The addition of this neutral salt is equivalent to a neutralization of the great bulk of the hydrochloric acid.

By adding an excess of sodium acetate, we can bring into play the common-ion effect described above and can thereby neutralize the acid still more.

As a proof of the reduction of the H^+-ion content of an acid solution by the addition of a salt of a weak acid, take the following illustration: A solution of ferrous ammonium sulphate is slightly acidified with hydrochloric acid, and then saturated with sulphuretted hydrogen. No precipitate of ferrous sulphide is obtained; on the addition to this

mixture, however, of a little concentrated sodium acetate solution, a black precipitate of ferrous sulphide is obtained.

In a similar manner, the concentration of OH' ions can be reduced by the addition of the salt of a weak alkali, e.g., by ammonium chloride.

10. The Neutralization of Acids and Bases.—The term "neutral" is often used rather vaguely, implying that a solution is neither acidic nor basic, as determined by its reaction to such an indicator as litmus; but if the investigator happens also to test the solution with some other indicator, he may get a different result, and therefore be confused. In other words, there are many substances which will react acid to one indicator and will give an alkaline reaction to some other indicator. It is evident, then, that one must have a clear understanding of the meaning of the term "neutrality" and of the nature of "indicators."

From the point of view of the ionic theory, an acid is any substance that can furnish H^+ ions, and a base any substance that can furnish OH^- ions, when dissolved in a dissociating solvent, such as water. Therefore, the term "neutrality" must refer to those cases where the concentration of both of these ions is the same, or where both are absent.

If HCl and NaOH are dissolved separately in water, and then mixed, it is quite common to express the reaction that takes place as follows:

$$NaOH + HCl = NaCl + H_2O.$$
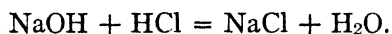
When, however, this reaction is considered from the standpoint of the ionic theory, we see that sodium hydroxide furnishes Na^+ and OH^- ions. When it is viewed from the standpoint of the law of mass action, it is evident that the following equilibrium is established:

$$
\begin{array}{c}
a \\
HCl \rightleftarrows H^+ \quad + Cl^- \\
b + \qquad + \\
NaOH \rightleftarrows OH^- \rightleftarrows Na^+ \\
\Updownarrow d \qquad \Updownarrow c \\
HOH + NaCl
\end{array}
$$
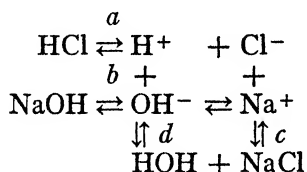
Since both HCl and NaOH are strongly dissociated, it is evident that when dilute solutions of these two substances are mixed, only a very few Na^+ and Cl^- ions will unite to form NaCl molecules (reaction c).

Water, on the other hand, is very slightly dissociated; hence practically all the H^+ and OH^- ions unite to form molecules of water (reaction d). The removal of these ions disturbs the equilibria and therefore enhances the reactions a and b, which, however, are almost complete. This means the almost complete cessation of any reunion of the ions Na^+ and OH^- or of Cl^- and H^+, so that in the end the solution will contain only NaOH, Na^+, and Cl^-, with a trace of undissociated NaCl molecules.

In brief, then, neutralization consists in the union of H^+ ions from an acid with OH^- ions from a base, to form the molecule water. The other ions remain in the solution mostly uncombined.

11. Hydrolysis.—There are few chemical reactions that are as important to the biochemist as the phenomenon of hydrolysis, which, as the name implies, is a breaking up, or decomposition, of compounds through the action of water, or, more accurately, by the ions of water, H^+ and OH^-. Nearly all the reactions of digestion and a great number of the reactions of metabolism are of this type. For example the hydrolysis of neutral fats or triglycerides into fatty acids and glycerol, of starches into sugars, and of proteins into amino acids belong to this category. At this time, however, we shall confine our attention to the discussion of the hydrolysis of salts.

Sodium carbonate in aqueous solution gives an alkaline reaction; ferric chloride in water gives an acid reaction; potassium cyanide in water gives an alkaline reaction and has the odor of hydrocyanic acid. Many other cases could be mentioned; in fact, in all aqueous solutions of salts of weak acids and bases we have the peculiar kind of electrolytic dissociation known as *hydrolytic dissociation*, i.e., the salt is decomposed by water with production of free acid and base:

$$AB + HOH \rightleftarrows HA + BOH \rightleftarrows H^+ + A^- + B^+ + OH^-.$$
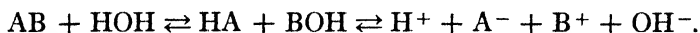
The reaction of this type of solution will now be either acid or alkaline, depending on the nature of the salt. Owing to the fact that salts are nearly always highly ionized, and that the anion or cation of the salt will combine with H^+ or OH^- ions from the water to form a nearly un-ionized acid or base, leaving an excess of OH^- or H^+ ions, respectively, in solution, the final result is an alkaline or acid reaction.

This behavior will be readily understood from a consideration of potassium cyanide. When this salt is dissolved in water, the following equilibria will indicate the course of the reaction:

$$KCN \rightleftharpoons K^+ + CN^-$$

$$+ \qquad +$$

$$HOH \rightleftharpoons OH^- + H^+$$

$$\Updownarrow$$

$$HCN$$
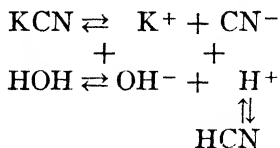
HCN is a very weak acid, being ionized in 0.1M solution only to the extent of 0.0001; hence, H^+ ions are removed from the solution by their union with CN^- ions. This upsets the other equilibria, necessitating the dissociation of more water and of more KCN. This increases the concentrations of K^+ ions and of OH^- ions and gives us, in effect, a solution of potassium hydroxide, which, being a strong alkali, remains almost entirely in an ionic state. We may, therefore, express this as a reaction:

$$K^+ + CN^- + OH^- = K^+ + OH^- + HCN.$$
$$H^+$$
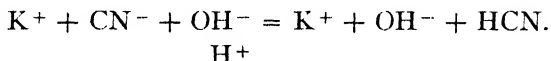
The acid reaction of ferric chloride is likewise just as easily explained. This salt combines with H^+ and OH^- ions to form ferric hydroxide and free hydrochloric acid. But the ferric hydroxide is a weak base, and hence is but slightly dissociated. The result is that OH^- ions are removed from the solution, thereby leaving free H^+ ions, which, of course, means that the solution will react acid. We may express this as a reaction:

$$Fe^{+++} + 3Cl^- + 3H^+ + 3OH^- = Fe(OH)_3 + 3H^+ + 3Cl^-.$$
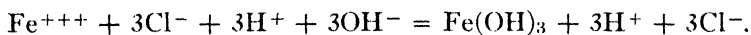
The reason why a salt of a strong base and a strong acid give a neutral reaction in aqueous solution may be seen from the following equilibria:

$$NaCl \rightleftharpoons Na^+ + Cl^-$$

$$HOH \rightleftharpoons OH^- + H^+$$
$$\Updownarrow \qquad \Updownarrow$$
$$NaOH \quad HCl$$
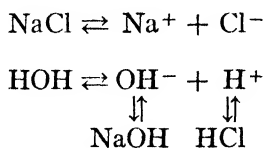
Two highly dissociated substances, NaOH and HCl, are produced, which means that neither H^+ ions nor OH^- ions will be tied up to give an unequal proportion, so that the reaction remains neutral.

The alkalinity of solution of sodium carbonate is due to the following hydrolytic reactions:

$$Na_2CO_3 \rightleftarrows 2Na^+ + CO_3^-$$

$$HOH \rightleftarrows H^+ + OH^-$$

$$HOH + CO_3^- \rightleftarrows HCO_3^- + OH^-$$

$$HOH + HCO_3^- \rightleftarrows H_2CO_3 + OH^-$$

$$H_2CO_3 \rightleftarrows CO_2 + HOH$$
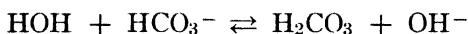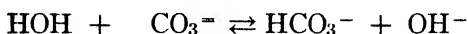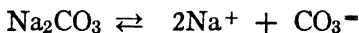
If the CO_2 is allowed to escape, the last reaction will only cease when all the H_2CO_3 has been decomposed. The final result is an excess of (OH^-) and, therefore, alkalinity.

In the case of the alkaline phosphates, the important hydrolytic equilibria will be represented by the expressions:

$$PO_4^{-} + HOH \rightleftarrows HPO_4^- + OH^-$$

$$HPO_4^- + HOH \rightleftarrows H_2PO_4^- + OH^-.$$
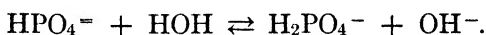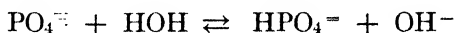
The salt NaH_2PO_4 has a slightly acid reaction in solution; Na_2HPO_4 has a slightly alkaline reaction; and Na_3PO_4 has a strongly alkaline reaction.

HPO_4^- is a very weak acid; hence PO_4^- occurs only in traces in solution.

12. Amphoteric Electrolytes and their " Iso-electric Points."—As was mentioned before (p. 174), an amphoteric electrolyte is a substance that can function as an acid by forming salts with bases, and as a base by forming salts with acids. It is evident that these substances must have two ionization constants, K_{H^+} and K_{OH^-}. It is understood, of course, that the concentrations of H^+ and of OH^- ions cannot both be very great in the same solution, since the equilibrium $(H^+) (OH^-) = K$ always holds (p. 184). Water has been described as an amphoteric electrolyte, since it dissociates both H^+ ions and OH^- ions. The dissociation constant for water at 22° C. is 1×10^{-7} for each of the ions. At this point the reaction is said to be neutral. Acids and alkalies upset the equilibrium between the H^+ and OH^- ions in water for a moment, but readjustment immediately takes place, and so the equilibrium is reëstablished and the value of K for $(H^+) (OH^-)$ will be 1×10^{-14}. In other words, if $(H^+) = 1 \times 10^{-5}$, then (OH^-) will be 1×10^{-9}.

An analogous condition holds for other amphoteric electrolytes. Let us take glycine, which is a typical amino acid and a typical ampho-

teric electrolyte. Glycine can form a sodium salt, CH_2—NH_2COONa, and also a hydrochloride, HCl—H_2N—CH_2—$COOH$. With strong acids it always behaves as a base; with strong alkalies, as an acid. In neutral solutions it is a feeble electrolyte, and is slightly dissociated into H^+ and an anion, thus:

$$H_2N\text{—}CH_2\text{—}COOH \rightleftarrows (H_2N\text{—}CH_2\text{—}COO^-) + (H^+) \qquad (1)$$
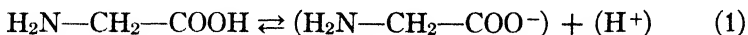
and partially into (OH^-) and a cation. Just as ammonia, when dissolved in water, exists, to a considerable extent, in the form of

$$\begin{array}{c} NH_3 \\ | \diagdown \\ H\ OH \end{array}$$

so it is supposed that glycine forms in water the compound

$$\begin{array}{c} CH_2\text{—}NH_2 \\ |\qquad |\diagdown \\ \qquad\ H\ OH \\ | \\ COOH \end{array}$$
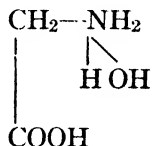
The equilibrium reaction may be represented as

$$HO\text{—}H_3N\text{—}CH_2\text{—}COOH \rightleftarrows (H_3{}^+N\text{—}CH_2\text{—}COOH) + (OH^+) \qquad (2)$$
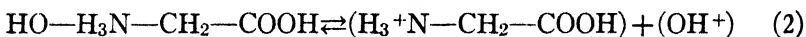
Reactions (1) and (2) may be combined in one equilibrium equation, thus:

$$\begin{array}{c} CH_2\text{—}NH_3\text{—}OH \\ | \\ COOH \end{array} \rightleftarrows \begin{array}{c} CH_2\text{—}NH_3{}^+ \\ | \\ COO^- \end{array} + (H^+)\cdot(OH^-) \qquad (3)$$
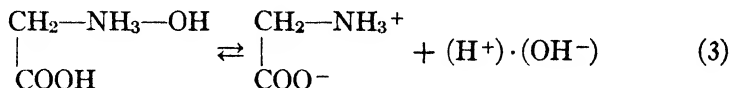
An ion of this type, which carries both a positive and a negative charge, is often called a "hermaphrodite" ion.

If a strong acid like HCl is added to an ampholyte (this is a shorter name for an amphoteric electrolyte) like glycine, a hydrochloride salt is formed. This affects a repression of the dissociation of the ampholyte into an anion and H^+ ions, in the same way that the dissociation of all weak acids is decreased by an increase in the H^+-ion concentration. On the other hand, the number of cations formed is increased, since the acid salt that is formed is highly dissociated. In the case of glycine we may represent the latter equilibrium as follows:

$$\begin{array}{c} C\text{—}CH_2\text{—}NH_2\text{—}HCl \\ | \\ COOH \end{array} \rightleftarrows \begin{array}{c} C\text{—}CH_2\text{—}NH_3{}^+ + Cl^- \\ | \\ COOH \end{array} \qquad (4)$$
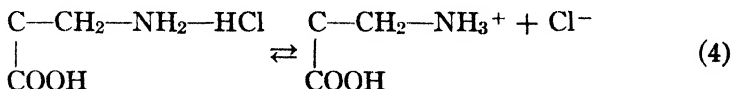
If a strong alkali, like NaOH, is added, the dissociation of the weak

base (2) is depressed, and there is an increase in the number of anions of the ampholyte, since the sodium salt formed is highly dissociated. The equilibrium may be represented as follows:

$$
\begin{array}{ccc}
CH_2\!\!-\!\!NH_2 & CH_2\!\!-\!\!NH_2 & \\
| & \rightleftarrows \quad | & +\ Na^+ \qquad (5) \\
COONa & COO^- &
\end{array}
$$
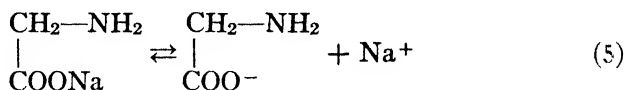
For every ampholyte there is a definite concentration of H^+ ions at which the electric conductivity is at a minimum. At this point, which is known as the "iso-electric point" for the ampholyte, either the substance will not migrate in an electric field, or half will go to the anode, and half to the cathode. Pauli has made an elaborate study of the iso-electric points of amino acids, and has made the following generalization: If we represent by K_a the dissociation constant of the carboxyl group of the amino acid, and by K_b the dissociation constant of the amino group, then when $K_a = K_b$ the iso-electric point will be reached in a neutral solution. If they do not equal each other, the amino acid will not be iso-electric in a neutral solution, but will be iso-electric at some other hydrogen-ion concentration. If the product of K_a and K_b is equal to 10^{-12} or more, there are no neutral molecules, but there are equal numbers (at the iso-electric point) of positively and negatively charged ions. If $K_a = K_b = 10^{-7}$ (their product $= 10^{-14}$), then there are a small number of neutral molecules at the iso-electric point, which is the point of neutrality, accompanied by equal numbers of positively and negatively charged ions. If $K_a = K_b = 10^{-8}$ (their product $= 10^{-16}$), then there are a large number of neutral molecules, other conditions being the same. If $K_a = K_b = 10^{-11}$ (their product $= 10^{-22}$), then all of the molecules are neutral at the point of neutrality of water, and there are no positively or negatively charged ions. Furthermore, the hydrogen-ion concentration may be varied between 10^{-6} and 10^{-8} without the appearance of an appreciable number of ions. In fact, it has to reach 10^{-3} in order to obtain 50 per cent of positively charged ions, or 10^{-11} to obtain 50 per cent of negatively charged ions. If K_a is not equal to K_b, the iso-electric point is not the point of neutrality of water, but the product of the relations is the same; i.e., in order to obtain the broad zone with all the molecules in neutral condition, it is necessary to have $K_a \times K_b = 10^{-22}$ or less.

It is known through the work of Emil Fischer that proteins are

built up from amino acids in peptide linkage, and, like the amino acids, they are also amphoteric electrolytes and show the same tendencies to combine with acids and alkalies to form true salts which undergo electrolytic dissociation. Whether the proteins combine with acids or with alkalies depends on the hydrogen-ion concentration. At a certain hydrogen-ion concentration which varies with each individual protein (according to the amino acids of which it is composed), the protein combines with neither acid nor alkali, and this particular hydrogen-ion concentration is the *iso-electric point* of the protein. At this point the protein is least soluble. It rarely happens that the iso-electric point of proteins is the same as the electrically neutral point of water; therefore, if we wish to precipitate proteins by the use of this principle, we almost invariably have to add a small amount of acid or alkali to the solution to bring it to the iso-electric point. In the case of most proteins, this point is reached by adding a small amount of acid.

The significance of the iso-electric point of various other substances, e.g., enzymes, will be considered later.

13. Solutions of Constant (H^+) and (OH^-): "Buffers."—We have already learned that the concentrations of (H^+) and (OH^-) in pure water at room temperature are almost exactly 1×10^{-7}. If small traces of certain electrolytes are added, this value is greatly changed. For example, one drop of normal acid or alkali added to a liter of pure water may increase this value two-hundred-fold. Now, it is very important to have solutions that have a constant (H^+) and (OH^-) value. This is especially true in the branch of analytical chemistry known as "micro-chemistry," which is so largely employed in the chemical analysis of blood and other biological fluids.

It has been found that sodium acetate not only decreases the (H^+) concentration of acetic acid but also *stabilizes* it; that is to say, it tends to protect it to some extent against alterations due to the addition of alkalies to the solution. It is clear that if we desire a solution to have the same reserve power for the neutralization of (OH^-) ions that it has for neutralizing (H^+) ions, we must have present the same concentration of acid ion and un-ionized acid. At these concentrations the value of the (H^+) ion will correspond to the ionization constant for the acid. This condition is observable for acetic acid in the equilibrium equation

$$\frac{(H^+) \times (ac^-)}{(H\,ac)} \rightleftharpoons 1.8 \times 10^{-5}.$$
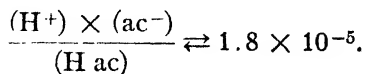
Let us now consider the effect of adding sodium acetate to a very dilute solution of hydrochloric acid. Both substances dissociate freely, and hence the following ions are originally present: Na^+, CH_3COO^-, H^+, and Cl^-. But, since acetic acid is a weak acid, CH_3COO^- and H^+ ions can exist together only in very low concentrations. Hence, we get the following equilibria:

$$HCl \rightleftarrows H^+ \qquad + Cl^-$$
$$+ \qquad\qquad +$$
$$CH_3COONa \rightleftarrows CH_3COO^- + Na^+$$
$$\downarrow \uparrow \qquad\qquad \Updownarrow$$
$$CH_3COOH \quad NaCl$$
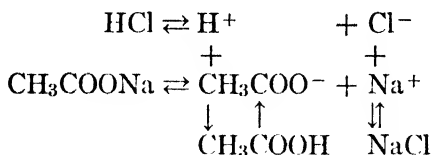
Stated in words, this means that the H^+ ions of the hydrochloric acid nearly disappear, owing to the presence of sodium acetate, and that the solution virtually becomes a solution of the weakly dissociable acetic acid. By controlling the concentrations of the HCl and CH_3COONa, we can produce a solution having a dissociation constant that is the same as that of acetic acid, i.e., 1.8×10^{-5}.

Substances that tend to hold the (H^+) and (OH^-) concentrations of a solution to a constant value are called "buffers." These substances play a very important rôle in nature and are widely used in the laboratory. We shall have occasion to consider their use in the next chapter in the preparation of standard solutions for the determination of hydrogen-ion concentration.

Many of the juices of the common fruits and vegetables have a rather high "total acidity" but a relatively low hydrogen-ion concentration, due to the presence of buffer salts.

In the animal body, proteins, amino acids, carbonates and phosphates all function as buffers. These help to maintain a fairly constant hydrogen-ion concentration in the tissue fluids.

In bacteriology, also, it is necessary to produce solutions of constant hydrogen-ion concentration for culture growth. For example, pure cultures of certain forms of yeast are cultivated in solutions buffered to a definite acidity.

CHAPTER X

MEASUREMENT OF HYDROGEN-ION CONCENTRATION

1. Importance of Hydrogen-ion Determination in Biological Processes.—It is very important to be able to ascertain with great exactness the true acidity or alkalinity of biological media. The old-fashioned way of stating that a fluid is acid or alkaline to litmus, etc., is not sufficient for the well-trained biologist or physician of to-day. Many biological processes are affected by minute changes in the reaction of the environment. In fact, the whole activity of an organism is largely regulated by reaction. Variations in acidity, for instance, are the causative factors in the regulation of respiration, the activity of muscle, and the excitability of nerve, and play a large part in regulating secretion and excretion. So important is the reaction of the body that nature has provided physical and chemical means to keep the normal reaction of the organism very close to the neutral point. Any marked deviation from this narrow reaction range is an indication of a pathological condition. We shall learn later on that the reaction of the organism is an equilibrium phenomenon, and that most metabolic disturbances have a tendency to upset this equilibrium.

As an example of the influence of small changes of reaction on the physiological behavior of a tissue, let us consider the heart. If the medium surrounding the heart, or the fluid within it, is too alkaline, it comes to rest usually in a contracted condition. On the other hand, if the fluid surrounding the heart is too acid, the beats cease, leaving the musculature in a completely relaxed condition. But if the fluid supplied is of just the right reaction, the heart will go on beating continuously, alternately relaxing and contracting. The important point to notice is that there is a very narrow range of reaction over which the heart muscle can function; outside the limits of this narrow range, its rhythmic activity cannot be maintained (Parsons).

Owing, then, to the exceptional importance, in all biological processes, of the acidity or alkalinity of the medium, it is necessary that

one be familiar with some of the methods by which the *actual* reaction, i.e., the concentration of free H^+ and OH^- ions, can be determined, and with the methods by which their concentrations can be regulated.

We have already learned that the concentrations of H^+ and OH^- ions are equal to a constant; therefore, it is evident that it is only necessary to determine one of these constituents. The other one can be readily calculated. As a general rule, then, the H^+-ion concentration is the only factor that is actually measured.

Several methods are available for the determination of H^+-ion concentration. They depend on (1) measurements of electromotive force of voltaic cells (electrometric methods); (2) the use of indicators; (3) the catalytic influence of hydrogen and hydroxyl ions on the velocity of certain chemical reactions. The third method depends on principles which we have not yet considered; hence it will not be described in the present chapter. The first two methods are the most important, or at least they are most frequently used by biochemists.

2. Review of the Principles Involved in the Determination of Hydrogen-ion Concentration.—(*a*) *Simple Acids.*—If we represent the electrolytic dissociation of an acid by the equation

$$HA \rightleftarrows H^+ + A^- \tag{1}$$

the equilibrium relations in very dilute aqueous solutions at constant temperature may be expressed by the mass action relation:

$$\frac{(H^+)(A^-)}{(HA)} = K_a, \tag{2}$$

where (H^+), (A^-), and (HA) represent molecular concentrations of hydrogen ions, anions, and undissociated molecules, respectively. (Review Chapters VIII and IX.) K_a is the dissociation constant. In general, the value of this constant does not vary much with moderate changes in temperature. From equation (2) it will be seen that if K_a is large there must be a relatively large proportion of the products of dissociation, and *vice versa* when K_a is small. The order of magnitude of the constant is an important characteristic property of each acid. In short, K_a is a measure of the "real strength" of an acid.

(*b*) *Polybasic Acids.*—Acids that contain more than one replaceable hydrogen ion per molecule ionize in what might be termed "steps,"

and each step has its own particular ionization constant. Thus, in the case of oxalic acid, the equilibria are as follows:

$$\frac{(H^+)(HC_2O_4^-)}{(H_2C_2O_4)} = K_{a,1} \tag{3}$$

and

$$\frac{(H^+)(C_2O_4^=)}{(HC_2O_4^-)} = K_{a,2} \tag{4}$$

Combining (3) and (4) we have

$$\frac{(H^+)^2(C_2O_4^=)}{(H_2C_2O_4)} = K_{a,1} \times K_{a,2} = K_a.$$

The value of $K_{a,1}$ has been found to be 1.0×10^{-1}, and that of $K_{a,2}$ to be 4.1×10^{-5} in the case of oxalic acid. It is typical of such acids that $K_{a,1} > K_{a,2} > K_{a,3}$, etc.

(c) *Bases.*—In a similar way, the dissociation of a base may be represented by the equilibrium equations

$$BOH \rightleftarrows B^+ \text{ and } OH^-, \tag{5}$$
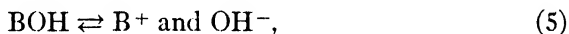
and

$$\frac{(B^+)(OH^-)}{(BOH)} = K_b. \tag{6}$$

Just as K_a is characteristic of an acid, so K_b is characteristic of a base. Likewise, just as the highly dissociated acids do not yield a constant K_a value, so the highly dissociated bases do not yield a constant K_b value.

(d) *Dissociation of Water.*—A most important relationship between acids and bases in aqueous solution is effected by the behavior of water, which dissociates to a slight extent into H^+ ions and OH^- ions, i.e., into the ion that is characteristic of acids and the ion that is characteristic of bases. This amphoteric action of water is extremely slight in magnitude, yet it is a factor of great importance in the numerous reactions which occur in aqueous solutions. The equilibrium of the reversible reaction $HOH \rightleftarrows H^+ + OH^-$ is represented by

$$\frac{(H^+)(OH^-)}{H_2O} = K. \tag{7}$$

Owing to the relatively enormous amount of undissociated water in relation to the dissociation products, the concentration of molecular water may be regarded as unchanged; hence K may be represented by another constant, and equation (7) may be written:

$$(H^+) \times (OH^-) = K_w). \tag{8}$$

At room temperatures—from 20° to 25° C. (average 22°)—$K_w = 1.0 \times 10^{-14}$ (approximately). This value is increased about fifty-fold at 100° C. It follows from this that even in the most acid solutions some OH^- ions remain, and *vice versa*. It further follows that in pure water, or in any neutral aqueous solution at room temperature, the value of (H^+) is the same as that of (OH^-), i.e.,

$$(II^+) = (OH^-) = 10^{-7}.$$

This means that water contains approximately 1 gram-molecule of $(H^+)(OH^-)$, or 1 gram of H^+ and 17 grams of OH^- ions in 10,000,000 liters.

It may be of interest to compare this value with that of decinormal HCl, that is, a hydrochloric acid solution containing one-tenth the molecular weight of HCl dissolved in enough water to make the total volume up to 1 liter. At this dilution, HCl is 91 per cent dissociated; therefore, the H^+-ion concentration (or C_H as it is written for short) is 0.091 N, or, in mathematical terms, $C_H = 9.1 \times 10^{-2}$.

3. Method of Expressing the H-ion (C_H) Concentration.—In order to avoid using several figures to express C_H, as has been done above, Sörensen,[1] introduced a scheme in which only one figure is required. This figure, designated by P_H, is found by subtracting from the power of 10 the common logarithm of the figure expressing the normality of the acid. Otherwise expressed, the P_H is the logarithm, to the base of 10, of the normality of the solution, with the minus sign omitted (but understood). Sörensen called this number the "hydrogen-ion exponent," and represented it by P_H. For typographical reasons it is now usually written pH.

There is still another way of expressing the pH of a solution: it is the logarithm of the number of liters that must be taken to contain 1 gram-ion of hydrogen. In the case of pure water at 22° C., this is exactly 10,000,000 liters. Now, log 10,000,000 = 7.

[1] Biochem. Zeit., **21**, 131, 1909.

It is evident that the concentration of hydrogen ions in 1 liter of water is 0.0000001 gram-equivalent per liter, or 1×10^{-7} normal. From these data we can deduce the following general relation:

$$p\mathrm{H} = \log_{10}\frac{1}{(\mathrm{H}^+)} \quad \text{or} \quad \log\frac{1}{(\mathrm{H}^+)} \tag{9}$$

and $(\mathrm{H}^+) = 10^{-p\mathrm{H}}$.

We have seen that $K_w = 10^{-14}$ for pure water at room temperature. If, therefore, in an alkaline solution at room temperature the concentration of OH ions is 0.001 normal (10^{-3}),

$$(\mathrm{H}^+) = \frac{K_w}{(\mathrm{OH}^-)} = \frac{10^{-14}}{10^{-3}} = 10^{-11}\mathrm{N}.$$

The following table will serve still further to elucidate the relationship:

TABLE XVII

RELATIONSHIP BETWEEN pH AND (H^+) VALUES

OH⁻ concentration....	10^{-13}	10^{-11}	10^{-9}	10^{-7}	10^{-5}	10^{-1}
H⁺ concentration....	10^{-1}	10^{-3}	10^{-5}	10^{-7}	10^{-9}	10^{-13}
pH................	1	3	5	7	9	13

It was previously stated that the pH of an acid is one of its characteristic properties; let us consider the pH of 0.1 N acetic acid, which is dissociated only to the extent of 1.36 per cent; C_{H} is therefore 0.00136 N, or 1.36×10^{-3}. The logarithm of $1.36 \times 10^{-3} = \log 1.36 + \log 10^{-3} = 0.13 + (-3.00) = -2.87$. This, with the minus sign omitted, is the so-called hydrogen-ion exponent, or pH, of 0.1 N acetic acid.

Let us take for another example the pH of 0.1 N NaOH. A 0.1 N solution of NaOH is 84 per cent dissociated at 20° C.; therefore, the OH⁻ ion is 0.084 N (i.e., 0.084 gram-equivalents of OH⁻ per liter). The product of H⁺ and OH⁻ concentrations is $10^{-14.14}$ at 20° C. Therefore, to get the pH of this solution we must get the logarithmic value of pOH to the base of 10 and subtract it from $10^{-14.14}$. $0.084 = 8.4 \times 10^{-2}$; the logarithm of 8.4 is 0.924, and this subtracted from the power $-2 = 1.08$ as pOH, or $14.14 - 1.08 = 13.06$, which is the value of the pH of the solution.

4. Unfitness of the Titration Method for Determining the Momentary Acidity i.e., the *p*H of a Solution.—It is important to note that the *p*H of a solution cannot be determined by the ordinary method of titration with an alkali of known strength. To illustrate this fact, let us consider the case of 0.1 N HCl and 0.1 N $H_4C_2O_2$. If these are titrated with 0.1 N NaOH solution until they each give a pink color with the *indicator*, phenolphthalein, 10 c.c. of each will require exactly 10 c.c. of the alkali to effect the change, and will therefore have the same apparent acidity. This is usually expressed by saying that equal volumes of 0.1 N hydrochloric acid and 0.1 N acetic acid have the same *total acidity*.

The total acidity of a solution is an expression of the total amount of replaceable hydrogen atoms present, that is, the free hydrogen ions plus those that will become free during the process of titration.

The amount of free hydrogen ions, or *p*H concentration, of the two solutions in question, however, is very different, being nearly seventy times as great in 0.1 N hydrochloric acid as it is in 0.1 N acetic acid. The reason for this, of course, is that hydrochloric acid is 91 per cent dissociated at 20° C., while acetic acid is only 1.36 per cent dissociated. If one neutralizes the solutions with a strong alkali, such as sodium hydroxide, one gets no indication of this difference in degree of dissociation, for as soon as the ionized part has been removed by the addition of the base, a further fraction of the previously undissociated acid is ionized. This process is repeated with further additions of alkali until all of the dissociable hydrogen ions have become dissociated and united with the OH$^-$ ions of the base. The next trace of alkali reacts with the indicator to give a pink color. (For the reason why an indicator changes color with change of reaction, see p. 217.)

Titration, therefore, only gives an index of the *capacity* of the solution to dissociate hydrogen ions; it does not give us any information concerning the *potential* of the hydrogen ions, i.e., the *p*H concentration. (This last statement shows the origin of the term *p*H or P_H, as Sörensen termed it.)

There are, however, methods by which the hydrogen-ion potential can be estimated without the liberation of any of the non-dissociated dissociable hydrogen ions. These will be considered below, under the headings of *electrometric methods* and *special indicator methods*.

5. Electrometric Methods for the Determination of Hydrogen-ion Concentration.—(*a*) *Electrode Potential.*—From the data presented in

Chapter VII on diffusion and osmotic pressure, it is evident that when a solid, such as sugar, is brought in contact with a solvent, such as water, it tends to pass into solution, that is, it disseminates itself throughout the liquid in such a manner as to suggest that it is under a definite pressure, and, from the analogy of this pressure to vapor pressure, we say that sugar has a definite *solution pressure*, which is constant at constant temperature. It was also made clear that the pressure of sugar in solution is its osmotic pressure. It was further shown in Chapter VIII that electrolytes have a solution pressure, which is often much greater than that of non-electrolytes like sugar.

When a metal is placed in contact with a solution of one of its salts, the metallic atoms tend to go into solution and become metallic ions; at the same time, this tendency is opposed by the osmotic pressure of the ions that are already present in the solution. The force with which the metal sends out its ions into the solution is known as the *electrolytic solution pressure*. If this pressure is equal to the osmotic pressure of the metallic ions in the solution, there is no difference of potential between the metal and the solution, and under such conditions there is no electric current generated. If the electrolytic solution pressure is greater or less than the osmotic pressure of the ions in solution, an electric current is established. When the electrolytic solution pressure is greater than the osmotic pressure of the ions, some of the metallic atoms will become ions by giving up electrons to the metal, thereby making the metal electrically negative; on the contrary, when the osmotic pressure is greater than the solution pressure, the positive ions from the solution deposit on the metal until the electrostatic changes stop further action. The metal is then positively charged and the solution negatively charged.

It has been found that the specific solution pressure of a metal is one of its characteristic properties. It is much greater in some cases than in others; for instance, the solution pressures of the alkali metals, zinc, iron, etc., are so great that they always exceed the osmotic pressures of their respective salt solutions, hence these metals are always negatively charged with reference to their solutions; on the contrary, the electrolytic solution pressures of mercury, silver, copper, etc., are so small that they become positively charged even in very dilute solutions of their respective salts. Electrolytic solution pressure varies with the temperature, with the nature of the solvent, and likewise with the concentration of the active substance in the electrode.

From the foregoing considerations, it is possible to derive an equation expressing the difference of potential between a metallic electrode and a solution of one of its salts. The exact relationship between the potential difference, E, the solution pressure, P, of the metal, and the osmotic pressure, p, of the metallic ions in the solution is established by thermodynamics, but the proof cannot be given here; the formula will be stated, however, and its significance explained. Briefly, then, provided that the changes at the junction of an electrode with a solution are reversible, the electromotive force at the junction can easily be calculated in terms of the solution pressure of the metal and the osmotic pressure of the solution. This can be done by calculating the maximum work obtainable when a gram-molecule of the metal electrode is brought from the pressure P to the lower pressure p, both osmotically and electrically. If a gram-molecule of a salt is brought reversibly from the pressure P to the osmotic pressure p, the osmotic work done may be expressed by the formula

$$\text{Osmotic work} = RT \log_e \frac{P}{p}. \tag{10}$$

The corresponding electrical energy gained is nFE, where E is the difference of potential between the metal and the solution, F is 1 faraday (96,500 coulombs), and n is the valence of the metal. Now the osmotic work done is equivalent to the electrical energy gained; hence, we may equate the two expressions as follows:

$$nFE = RT \log_e \frac{P}{p}, \tag{11}$$

or

$$E = \frac{RT}{nF} \log_e \frac{P}{p}. \tag{12}$$

In order to obtain E in volts, R must be expressed in electrical units (volt-coulombs). Now, $R = 1.99$ cals., and the electrical equivalent of heat $= \dfrac{1}{0.2394} = 4.18$; therefore, $R = 8.32$ in electrical units.

Ordinary, or natural, logarithms are reduced to logarithms to the base 10 by multiplying by the factor 2.303.

T at $18°$ C. $= 273 + 18 = 291$.

If we now substitute the foregoing figures in the equation

$$E = \frac{RT}{nF} \log_e \frac{P}{p},$$

we get, when $n = 1$,

$$E = \frac{8.32 \times 291 \times 2.303}{96,540} = 0.058 \cdot \log_{10} \frac{P}{p}. \qquad (13)$$

The value of the expression is taken as $0.058 \log_{10} \frac{P}{p}$ at room temperature for monovalent ions, and the general formula becomes

$$E = \frac{0.058}{n} \log_{10} \frac{P}{p}, \qquad (14)$$

which should be remembered.

From the above data it is clear that a ten-fold change in osmotic pressure gives a change of 0.058 volt in the potential. (The solution pressure of a metal remains constant, so that any change in potential is due entirely to change in osmotic pressure.) This means that every time we have a ten-fold change in concentration of a monovalent ion we have a ten-fold change in osmotic pressure, which consequently must change the potential by 0.058 volt. For a divalent ion, a ten-fold change in concentration will give only half as much change in osmotic pressure, and, therefore, will change the potential only 0.058/2 volt.

The relationship between changes in osmotic pressure and changes in electrical potential is a very important phenomenon from the standpoint of the physiologist or physician, and therefore should be carefully noted by the student.

(b) *A Primary Cell.*—A good illustration of the part played by the osmotic pressure of metallic ions in determining the potential difference at a metal electrode is obtained from the familiar Daniell cell, which consists of a zinc rod immersed in a solution of zinc sulphate separated by a porous pot from a solution of copper sulphate in which a copper electrode is placed; when the zinc and copper poles are connected externally with a metal wire, a current flows through the wire from the copper to the zinc. In the Daniell cell, usually represented as $Zn/ZnSO_4$, Cu/SO_4, there are, as in every other simple galvanic cell, two places where the junction metal/solution occurs, and the potential

difference at each of the two electrodes can be expressed by the formula

$$E = \frac{0.058}{n} \log_{10} \frac{P}{p}.$$

Figure 19 will serve to indicate the relative positions of the several components of either of these cells:
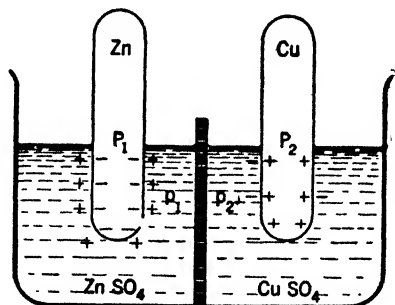
FIG. 19.—The Daniell Cell.

The double layer is formed on both. Until the metal electrodes are connected by wire, the equilibrium will be undisturbed. As soon as they are connected, a current will flow from copper to zinc, i.e., Zn ions will go into the solution, depositing an equivalent amount of Cu ions upon the copper plate. The electromotive force will then depend on the difference between the two values of P, p being the same on both sides, and the amount of current will depend upon the number of gram-equivalents of Zn dissolved and of Cu precipitated.

If we ignore the slight potential difference which arises at the common surface of the two solutions, we can express the E.M.F. of the Daniell cell as $E_1 - E_2$, where E_1 and E_2 are the potential differences at the Zn and Cu electrodes, respectively. Now if P_1 and P_2 represent the electrolytic solution pressures of Zn and Cu, and p_1 and p_2 represent the osmotic pressures of Zn ions and Cu ions, respectively, in the two solutions, then

$$E_1 - E_2 = \frac{0.058}{2} \log_{10} \frac{P_1}{p_1} - \frac{0.058}{2} \log_{10} \frac{P_2}{p_2} = \frac{0.058}{2} \log_{10} \frac{P_1 p_2}{P_2 p_1}. \quad (15)$$

This final formula is the E.M.F. of the Daniell cell. One can readily draw the conclusion from this formula that by diluting the zinc sul-

phate solution we diminish the value of p_1, and by concentrating the copper sulphate solution we increase the value of p_2. The effect is the same: the E.M.F. of the Daniell cell is raised.

(c) *Concentration Cells.*—We have now to consider what are termed "concentration cells," i.e., cells in which the E.M.F. depends essentially on differences of concentration. From the data presented regarding the relationship between osmotic pressure and electrolytic potential, it is evident that it should be possible to construct a cell having the same metal and the same solution at both electrodes, simply by making the solution more concentrated at one side. The arrangement would be similar to the Daniell cell, but we should have one metal, say zinc, and one solution, say zinc sulphate, at both sides. In such a cell, a counter E.M.F. develops at each electrode; therefore, if the two solutions are of equal concentration, they will neutralize each other, and no current will flow. But if one solution is much more concentrated than the other, the E.M.F. on one side will be greater than on the other, and a current will flow, provided suitable connections are made.

In the case of concentration cells also, the potential can be calculated from the difference in concentration and the valence of the ion, by the use of the fundamental equation. In this case C_1 is substituted for P_1, where C_1 represents an ionic concentration that would just balance the solution pressure P_1; in like manner, C_2 is substituted for P_2. For p_1 and p_2 the corresponding ionic concentrations may be substituted by the terms c_1 and c_2. The equation then becomes

$$E = E_1 - E_2 = \frac{0.058}{n}\left(\log_{10}\frac{C_1}{c_1} - \log_{10}\frac{C_2}{c_2}\right). \qquad (16)$$

But, since the metal is the same at both electrodes, $C_1 = C_2$; the equation may be simplified to

$$E = E_1 - E_2 = \frac{0.058}{n}\left(\log_{10}\frac{c_2}{c_1}\right); \qquad (17)$$

and, in the case of a simple half-element,

$$E_1 = E_2 + \frac{0.058}{n}\left(\log_{10}\frac{c_2}{c_1}\right), \qquad (18)$$

where E_1 is the electrode potential for ionic concentration c_1, and E_2 = ionic concentration for c_2. If $c_1 = 1$, the equation becomes

$$E_1 = E_2 + \frac{0.058}{n} \log_{10} c_2. \tag{19}$$

(*d*) *Hydrogen Electrode.*—We may now proceed to the concentration of the *hydrogen electrode*. From the foregoing discussion of cells, one might easily infer that the electrodes are invariably of metal; the principle has been extended, however, to cover cases where the electrode substance is really a gas. Suppose, for example, that in place of using a metal in the construction of one of the electrodes, we use an electrode consisting of a layer of pure hydrogen gas in contact with a solution containing free hydrogen ions; then the rate at which hydrogen ions become added to the solution from the H_2 layer, or taken from it, will depend on the concentration of H^+ ions in solution. It is evident that hydrogen gas, as such, cannot be used, and that a means must be found which will furnish a hydrogen layer. This can fortunately be accomplished by making use of the property which certain metals, such as palladium and platinum, have of absorbing large quantities of hydrogen gas.

A strip of platinum, covered over electrolytically with platinum black, absorbs on its surface a large quantity of hydrogen gas, and hence may be made to serve as a hydrogen electrode. If such a strip is partially immersed in a solution containing hydrogen ions (say dilute HCl), there is exhibited a difference of electrode potential between the solution and the electrode, which is dependent upon the concentration of H^+ ions, just as the potential difference between a zinc electrode and a solution containing zinc ions is dependent upon the concentration of the zinc ions. The hydrogen electrode prepared in this manner behaves, then, just as if it were a sheet of metallic hydrogen. If positive electricity passes through the solution, H^+ ions are discharged then $2H^+ = H_2$; if negative electricity passes, the gaseous hydrogen becomes ionized then $H_2 = 2H^+$.

An oxygen or chlorine electrode behaves in a similar manner. In the case of oxygen, the platinum strip saturated with O_2 is immersed in an alkali solution, which may be regarded as containing oxygen ions, derived from the OH^- ions, thus: $2OH^- \rightleftarrows O^= + H_2O$.

The vessel used in the construction of a "gas electrode" is similar in principle to the one used with metal electrodes, but has an additional

feature in that it is so constructed as to permit a current of gas to be bubbled through the solution. The arrangement of the apparatus can be seen from Fig. 20.

It should be evident that the E.M.F. of a hydrogen "gas-chain" electrode can be determined in a similar manner to that described for a concentration cell. In short, the E.M.F. is determined solely by the relative concentration of hydrogen ions around the two electrodes. If C_1 represents the concentration of the hydrogen ions in one solution, and C_2 the concentration in the other, the E.M.F. of the combination will be obtained at room temperature by the formula

$$E = 0.058 \log_{10}\frac{C_1}{C_2},\qquad(20)$$

provided we disregard the interference which occurs at L. This will be discussed below.
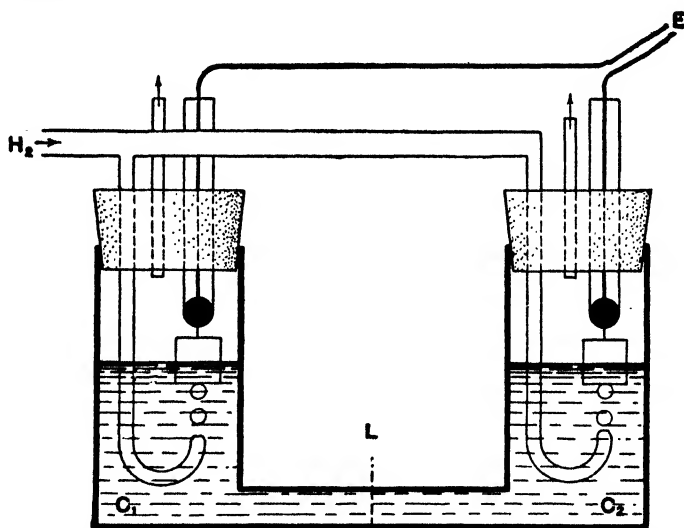
FIG. 20.—Concentration Chain of Hydrogen Electrodes. (After Clark.)

If, then, we know one concentration and determine the ratio of the two from the E.M.F. by means of the above formula, the concentration of the other can be calculated. If, for example, $C_1 = 1$ gram-equivalent of H^+ ion per liter, and if we represent C_2 by (H^+), we can obtain the value of the unknown hydrogen-ion concentration by means of the expression

$$E = 0.058 \log_{10}\frac{1}{(H^+)}.\qquad(21)$$

As stated above, the expression $E = 0.058 \log_{10} \dfrac{C_1}{C_2}$, in obtaining the E.M.F. of a hydrogen-concentration cell of the type just described, is often subject to a fairly large correction. This is due to the fact that differences of potential may be established at the junction of the two acid solutions, as well as at the electrodes, owing to the greater speed with which hydrogen diffuses. From the table on page 171 we learn that hydrogen ions travel about five times as fast as chlorine ions. This means that in a cell of the type shown in Fig. 20 there will be a difference of potentital at L, since there will be an unequal distribution of H^+ and Cl^- ions in the two cells, due to the different speeds with which they travel. Ultimately, of course, when the two solutions
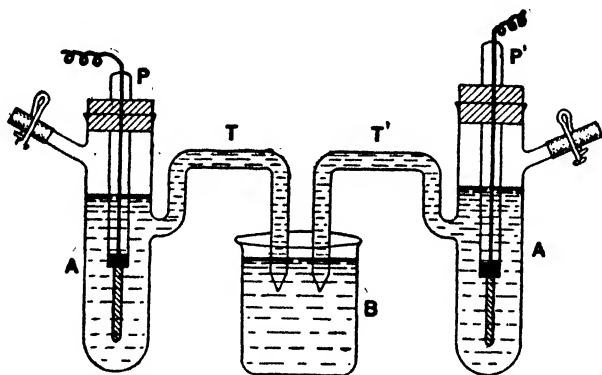
Fig. 21.—Hydrogen Electrode Showing KCl Connection Vessel.

have acquired the same concentration, the difference of potential will cease to exist. Nernst[2] has evolved a formula which corrects these differences of potential, but since it is rather complicated it will not be discussed here. Besides, in the case of complex physiological fluids, such as blood, calculations by the use of the Nernst formula are practically impossible, since the actual ions present are an uncertain quantity. There is another simple way of overcoming the difficulty, namely, by inserting a saturated solution of potassium chloride between the two acid solutions. The manner in which this is done can be seen from Fig. 21. The electrode P is in contact with an acid solution in cell A which contains $(H^+) = C_1$; another electrode, P', is in contact with an acid solution in cell A' which contains $(H^+) = C_2$.

[2] Zeit. phys. Chem., **4**, 129, 1889.

The acid solutions are connected by means of the tubes T and T' with vessel B which contains the satu_ed solution of KCl. The introduction of KCl into the system reduces the difference of potential to a minimum, so that in this case the E.M.F. of the cell is equal to the difference in the potential of the two electrodes, i.e., $E = 0.058 \log_{10}\dfrac{C_1}{C_2}$.

(c) *The Calomel Electrode.*—It has been found that it is unnecessary for both of the electrodes to be hydrogen electrodes, or to have a connection battery in hydrogen. So long as the opposing electrode is of known electromotive force, it may be of any form. In other words, in practice it is desirable to have a reliable standard electrode, and since hydrogen electrodes are liable to undergo change after they have been standardized, it is customary to use an *Ostwald calomel electrode* as the standard electrode. This electrode consists of mercury in contact with a solution of normal KCl saturated with pure mercurous chloride. The principal advantage of this electrode is that it can be reproduced with an accuracy of about 1 millivolt. A common arrangement is seen in Fig. 22. At the bottom of the container there is a layer, of mercury
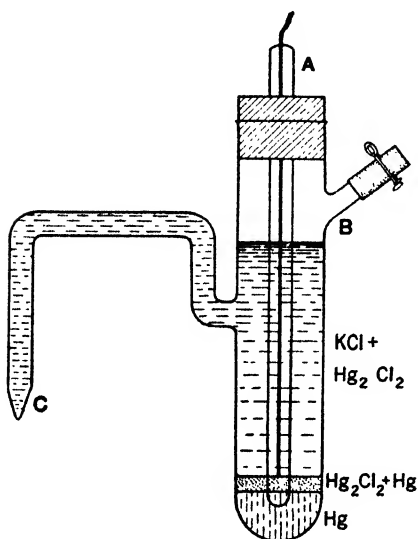
Fig. 22.—Calomel Electrode.

then a paste made by rubbing, in a mortar, mercury and calomel plus some of the KCl solution, and then the vessel is filled up with N KCl solution which has been previously saturated with calomel by shaking with excess of the latter. Electrical connection is obtained through the platinum wire A, which passes through the bottom of the vertical tube to the mercury, and through the tube C, which must be filled with KCl solution.

Under normal conditions a potential of $+ 0.56$ volt is developed in the calomel electrode. The positive sign before the value 0.56 indicates that the electrode is positive in relation to the solution.

In the use of calomel electrodes in pH determinations, however, it

is customary to make actual measurements with the calomel electrode, and then refer them to the hydrogen standard, taking the potential of the calomel electrode to be $+ 0.283$ volt, when referred to the normal hydrogen electrode as zero.

(*f*) *Measurement of Hydrogen-ion Concentration with the Calomel Electrode.*— The hydrogen-ion concentration of a solution can readily be calculated from the measured value of the electromotive force of a calomel electrode cell of the type just described. An idea of the arrangement of the complete apparatus in which the calomel electrode is used can be obtained by examining the accompanying sketch (Fig. 23).
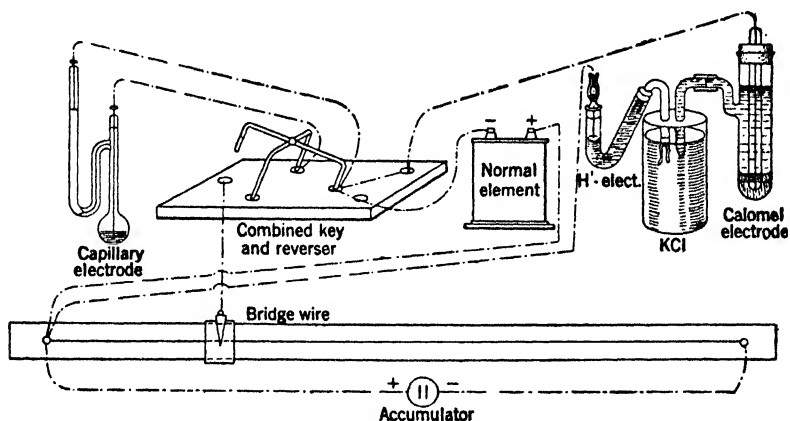
FIG. 23.—Apparatus for Measuring the Hydrogen-ion Concentration.

The current that is generated in the battery, which is composed of the calomel electrode, connecting vessel with KCl solution, and the hydrogen electrode, or that from the normal element, is transmitted through a reversing key to the bridge wire, where the voltage is compared with a steady current flowing through the bridge wire from an accumulator. The capillary electrometer is used to detect the flow of the current at various positions of the movable contact on the bridge wire.

A description of the technique for measuring the hydrogen-ion concentration would be out of place here, especially as each make of apparatus has its own distinctive features. The principle of the method is as follows: A hydrogen electrode is immersed in the solution to be tested, and the E.M.F. of the cell, formed by connecting the solution

with the calomel electrode, is measured. This value is then employed in the equation given below, and the value of pH is then calculated.

For example, let us suppose the measured E.M.F. to be 0.667 volt; the value of the hydrogen-ion concentration would then be calculated as follows: The E.M.F. (measured) − 0.283 (for origin of this figure see p. 212) $= 0.058 \log_{10} \dfrac{1}{(H^+)}$.

or,

$$\log_{10} \frac{1}{(H^+)} = \frac{\text{measured E.M.F.} - 0.283}{0.058} = p\text{H}.$$

Substituting, we have

$$\log_{10} \frac{1}{(H^+)} = \frac{0.667 - 0.283}{0.058} = 6.6206.$$

$$\therefore \ \log (H^+) = 0.0000 - 6.6206 = \bar{7}.3794 \left(\text{The co-log of } \frac{1}{(H^+)}\right).$$

From this $(H^+) = 2.396 \times 10^{-7}$.

(g) *Determination of Hydrogen-ion Concentration by Means of a Potentiometer.*—One of the most practicable methods of determining the pH of a biological fluid is by means of a "potentiometer," which consists of a wire through which an electric current is flowing. Along this wire there is a drop in potential, so that by determining the fall of potential it is possible to pick out two points of any difference of potential. By using a wire of uniform thickness, and therefore of uniform resistance, it is possible to secure a uniform drop. It is customary to make this wire 1.083 meters in length, and to pass a current through it from an ordinary lead cell. A rheostat is placed in the circuit in order that the amount of current and difference in potential of the two ends of the potentiometer wire may be varied. These two ends are connected by means of a key with a standard Weston cell which has an E.M.F. of 1.083 volts. By placing a galvanometer in the circuit with the standard cell, one can ascertain whether or not a current is flowing. If we now adjust the rheostat so that the two ends of the potentiometer wire have a difference in potential of 1.083 volts and the poles of the standard cell are connected in the same direction as the poles of the potentiometer wire, no current will flow through the galvanometer. This condition is obtained when each millimeter in the potentiometer wire shows a difference in potential of 1 millivolt. When

ᴀg a determination of the E.M.F. of a hydrogen-calomel cell, the ᵣ is turned so that the calomel electrode is marked + and the ᴧydrogen −, and set in place of the Weston cell. The contact from the positive pole is then made to slide along the potentiometer wire until the galvanometer shows no deflection. The number of millimeters from the negative end to the sliding contact is a direct index of the millivolts of the E.M.F. of the hydrogen-calomel electrode cell. The *p*H is now determined by the use of a standard formula.

There are also direct-reading potentiometers, especially adapted for the use of biologists and bacteriologists, who have to determine the acidity and alkalinity of many physiological fluids that are so highly colored that the standard colorimetric methods (described below) are not practicable. (See article by Dr. W. T. Bovie in the *Journal of Medical Research*, Vol. 33, p. 295, November, 1915.) In the Bovie Hydrogen-ion Potentiometer the concentrations of the H^+ and OH^- ions can be read directly by means of a special scale. The measurement is made by rotating the slide wire until the galvanometer deflection is zero, which indicates that the electromotive force of the potentiometer and measuring circuit exactly balance the potential difference between the hydrogen electrode and the calomel electrode. As the slide wire rotates, a marking device connected with it moves across a chart on which the concentration scale is ruled. When the galvanometer deflection is zero, the marking device is pressed against the chart. This will print a dot enclosed in a circle, which will show the concentration of the hydrogen ions in the solution. The numbers on the ordinate ruling of the chart are exponents of the negative power of 10; therefore, the readings indicated are the *p*H of the solution. Figure 24 shows the complete Bovie apparatus. (NOTE.—These instruments can be obtained from Eimer and Amend, New York, and a complete description of the apparatus with directions for using it is furnished by them.)

6. The Indicator Method of Determining Hydrogen-ion Concentration.—The electrometric method for the determination of the hydrogen-ion concentration of liquids is the most accurate means that has yet been found for this purpose. It is, therefore, the standard method against which all other methods are checked. But it involves expensive and intricate apparatus, and hence is not universally used, especially since a simple and less expensive method has been devised, which is quite accurate and easily performed after a little practice.

The author refers to the so-called "indicator," or "colorimetric," method. This method is now quite generally used in the routine determinations of hydrogen-ion concentrations in biological and medical laboratories.

The principle of the method depends on the fact that there are certain dyes, or "indicators," which change in color tone or in depth of color with changes in the pH of a solution. The method must not be confounded with the "titration" method in which indicators are also used. It has already been pointed out on page 203 that the titration method tells us the total acidity or alkalinity of a solution, but not the
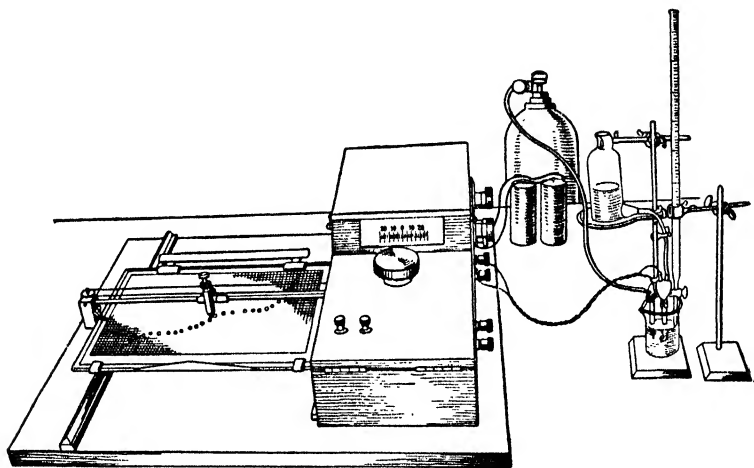
FIG. 24.—The Bovie Potentiometer.

hydrogen-ion concentration. In the titration method the important factor is the accuracy of the standard acid or alkali used for neutralization. In this case the indicator just serves to show the change of reaction. In the colorimetric method the indicators play a far more important part; a large number of indicators are used and therefore it is desirable, for one's understanding of the method, to have a very clear idea of the nature of these substances. Although, as a matter of fact, one can use the method without this knowledge, yet better work can undoubtedly be done with it. To quote Clark:[3] "In the first place it should be emphasized that the customary manner of using indicators for the determination of hydrogen-ion concentration is

[3] The Determination of Hydrogen Ions, Williams and Wilkins, Baltimore, 1920.

basically a comparative method with hydrogen electrode values as 'calibration.' With standard buffer solutions whose hydrogen-ion concentrations have been determined, indicators may be arranged empirically without involving any theory whatsoever. It is well to emphasize this uninspiring, matter-of-fact aspect of the matter, because it will remind us that, if so much of practical value has been done without the aid of theory, the application of theory may lead on to greater things."

Theory of Indicators.—Indicators are weak organic acids or bases (mostly acids) which possess one color in an acid solution and a different one in an alkaline solution. Ostwald[4] advanced the theory that indicators are very weak acids or bases whose undissociated molecules have a different color from that of their dissociated molecules. Phenolphthalein, for example, is a very weak acid, and therefore, in conformity with the theory of electrolytic dissociation, may exist almost entirely in the undissociated form in the presence of a very low concentration of hydrogen ions. In the presence of a base, a highly dissociable salt is formed, and the anion produced imparts a pink color to the solution. In the titration of a weak acid with NaOH solution, the successive reactions are as follows, when HPh represents the undissociated phenolphthalein, and Ph^- the anion:

$$(a)\ \ HA\ \ + Na^+ + OH^- = Na^+ + A^- + HOH;$$

$$(b)\ \ HPh + Na^+ + OH^- = Na^+ + Ph^- + HOH.$$
$$\quad\quad\ \ \text{Colorless} \quad\quad\quad\quad\quad\quad\quad\quad \text{Pink}$$
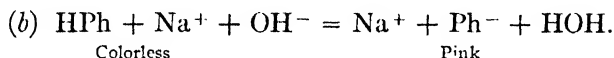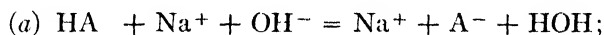
(a) is nearly complete before (b) commences, owing to the very low dissociation constant of $HPh(2 \times 10^{-10})$.

Phenolphthalein has been found to be an unsatisfactory indicator for weak bases such as ammonia. This is explained by the hydrolysis of the salt of the indicator, which prevents a visible amount of Ph^- from being formed until a considerable excess of NH_4OH has been added. This may be expressed thus:

$$NH_4^+ + Ph^- + HOH \rightleftarrows HPh + NH_4OH.$$

On the other hand, methyl orange is a weak base which changes in color at pH 4; it is therefore unsuitable for the titration of a weak

[4] Foundations of Analytical Chemistry, 124; 2nd ed., 1900. The Macmillan Co., New York.

acid, but is suitable for the titration of a weak base. This is likewise explainable on the grounds of the hydrolysis of the indicator salt.

One can titrate solutions of strong acids with a strong base, or *vice versa*, with any indicator whose color change occurs within the pH range 4–10, because hydrolysis is negligible.

Ostwald's theory has been found to agree with the observed facts in most cases, but certain anomalous results were unaccounted for, until it was shown that the process is not a simple dissociation phenomenon but that the color change is due to an intromolecular rearrangement. In other words, it is now known that most indicators are weak organic acids which are capable of existing in two tautomeric forms in equilibrium with each other. One form prevails in acid solutions and the other in basic solutions. There may be two or more color changes which occur at different C_H concentrations.

The tautomeric forms of phenolphthalein have been shown by Rosenstein[5] to be

$$C_6H_4 \overset{\displaystyle\hspace{-0.2em}}{\underset{\displaystyle \begin{matrix} | \\ C \\ \| \\ O \end{matrix}}{\hspace{-0.2em}}} \hspace{-0.4em}\begin{matrix} \\ \text{---}C \end{matrix}\!\!\begin{matrix} C_6H_4OH \\ \diagdown \\ | \quad C_6H_4OH \\ \text{---}O \end{matrix}$$
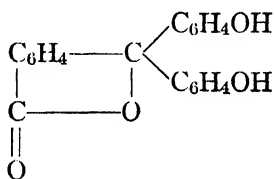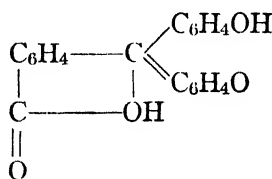
I. Colorless.   II. Pink.

The mass-action expressions for tautomeric equilibria have been formulated by Acree[6] and Noyes.[7] The following quotation is from the paper by Noyes:

"Let us consider an acid indicator composed of two tautomeric substances, HIn' and HIn'', which when neutralized with a base BOH are converted into the corresponding salts BIn' and BIn'', these being largely dissociated into the ions B', In'^- and BIn''^-. In order that an indicator may be satisfactory in the respect that it shows a sharp color-change on neutralization, it is evidently essential that the following conditions be fulfilled: (1) One of the structural forms must be colorless and the other colored; or else, the two structural forms must be of distinctly different colors. (2) Either form that may be colored must not exist in appreciable proportion in both acid and alkaline solutions.

[5] J. Am. Chem. Soc., **34**, 1117, 1912.
[6] Am. Chem. J., **38**, 1, 9107; *ibid.*, **39**, 529, 1908.
[7] J. Am. Chem. Soc., **32**, 815, 1910.

"In accordance with these general conditions, the three following cases may occur, the convention that the indicator shows in acid solution the color corresponding to the form HIn' being employed in their fomulation:

"(a) The indicator is colorless in acid and colored in alkaline solution. Hence, HIn', BIn', and In'⁻ are colorless; HIn'', BIn'', and In''⁻ are colored; and (HIn'')/(HIn')⁸ must be very small.

"(b) The indicator is colored in acid and colorless in alkaline solution. Hence HIn', BIn', and In'⁻ are colored; HIn'', BIn'', and In''⁻ are colorless; and (In'⁻)/(In''⁻) must be very small.

"(c) The indicator has two sharply differentiated colors in acid and alkaline solution. Hence HIn', BIn', and In'⁻ have one color; HIn'', BIn'', and In''⁻ have a different color; and both (HIn'')/(HIn') and (In'⁻)/(In''⁻) must be very small (otherwise a mixed color would be obtained).

"Corresponding criteria for a satisfactory color-change can also be readily formulated for the case of a basic indicator existing in alkaline solution in the two structural forms In'OH and In''OH, which when neutralized with an acid HA are converted into the corresponding salts In'A and In''A, these being largely dissociated into the ions In'+, In''+, and A⁻. It is not. however, necessary to formulate these conditions here.

"We may next derive a general expression (as has been previously done by Acree[9]) for the equilibrium-relations of any pair of tautomeric acids and their ions. The three fundamental equilibrium equations are as follows:

$$\frac{(H^+)(In'^-)}{(HIn')} = K_I'; \qquad (22)$$

$$\frac{(H^+)(In''^-)}{(HIn'')} = K_I''; \qquad (23)$$

and

$$\frac{(HIn'')}{(HIn')} = K_T. \qquad (24)$$

---

[8] NOTE.—Symbols enclosed within parentheses denote the molar concentrations of the substances represented by the symbols.

[9] *Loc. cit.*

Multiplying (23) by (24), adding (22) to the product, and substituting in the denominator for (HIn′) its value $\dfrac{(\text{HIn}') + (\text{HIn}'')}{1 + K_T}$ given by (24), we get

$$\frac{(\text{H}^+)[(\text{In}'^-) + (\text{In}''^-)]}{(\text{HIn}') + (\text{HIn}'')} = \frac{K_I' \times K_I''K_T}{1 + K_T} = K_{IA}. \qquad (25)$$

"If the indicator is a base existing as the two tautomeric substances In′OH and In″OH, having the ionization constants K_I' and K_I'' and a tautomer constant K_T defined by equations analogous to (22), (23), and (24), the general expression for the equilibrium between the ionized bases and their ions is

$$\frac{(\text{OH}^-)[(\text{In}'^+) + (\text{In}''^+)]}{(\text{In}'\text{OH}) + (\text{In}''\text{OH})} = \frac{K_I' + K_I''K_T}{1 + K_T} = K_{IB}." \qquad (26)$$

Noyes has used in the above expressions a single constant K_{IA} or K_{IB} in place of the function of the three constants K_I', K_I'' and K_T, to denote the ionization constant of the indicator acid or base. He calls K_{IA} and K_{IB} the "apparent dissociation constants." From an inspection of equations (25) and (26) it is evident that it is the degree of dissociation, as determined by the hydrogen-ion concentration, that determines which tautomer predominates. Consequently, the later work on tautomeric equilibria only modifies the original Ostwald dissociation theory.

Expression (25) may also be written in the following simplified form:

$$\frac{(\text{H}^+)(\text{In}^-)}{(\text{HIn})} = K_{IA}, \qquad (27)$$

where (HIn) denotes the sum of the concentrations of substances in the acidic form, and (In⁻) the sum of those in the basic form. If, now, we let b equal the concentration of the basic, and $(1 - b)$ the concentration of the acidic forms, and denote the degree of dissociation of the indicator salt by α_{IS}, then we may write

$$(\text{H}^+) = K_{IA}\frac{(\text{HIn})}{(\text{In}^-)} = K_{IA}\frac{(1 - b)}{\alpha_{IS}^b}, \qquad (28)$$

and for a basic indicator,

$$(H^+) = \frac{K_w}{K_{IB}} = K_I \frac{\alpha IS(1-b)}{b};$$ (29)

and if we take the degree of dissociation of the indicator salt as unity both the equations (28) and (29) may be written thus:

$$(H^+) = K_I \frac{(1-b)}{b}.$$ (30)

Perhaps the most important result of the above study of indicators is that it has been found that when the indicator is 50 per cent dissociated, i.e., when half of the indicator is transformed from the acidic to the basic form, we have the relation $K_I = (H^+)$. This relation is of great importance in determining the hydrogen-ion concentration of solutions by the "colorimetric" method (see below), or, conversely, in the determination of the "apparent dissociation constants" of indicators.

The precise pH at which an indicator is 50 per cent dissociated depends on the individual properties of the indicator, and on the nature of the solution which is being examined. (Foreign substances may interfere with the result.) The pH of a solution is 9.7 when phenolphthalein is 50 per cent dissociated, and it is 4.06 when 2–4 dinitrophenol is 50 per cent dissociated, etc.

From the standpoint of the general use of indicators, however, the value of the so-called "end point" or "turning point" of an indicator is more important than that of the constant. This is evident, since one titrates to the point at which one observes most clearly a definite color change, rather than to the point of 50 per cent dissociation. To obtain this latter value, one has to use a colorimeter and observe many other troublesome precautions. Moreover, there are a number of indicators whose end point can be obtained (after a little training of the eye) at a pH not far from the "half-transformation point."

In the use of indicators there are several precautions which must be observed. For instance, the solution should be kept as free as possible from foreign substances, since the hydrogen-ion concentration at which certain indicators change color is not the same in pure acids or bases as in the presence of these substances. Salts and proteins are especially troublesome. Likewise, it is evident that one must not use

too much indicator, firstly, because an excess of the indicator will obscure the end point, and, secondly, since the indicators are weak acids or bases, and it will therefore be obvious that an appreciable neutralization effect might take place if an excess of indicator were used.

The table given in Fig. 25, which is extracted from the results of Salm (1906), Sörensen (1909), and Bayliss (1914), contains a carefully selected group of indicators. These indicators have been found to be unaffected by the presence of moderate amounts of such neutral salts or proteins as are likely to be present in physiological solutions.

The double vertical line implies that the change of color is sharpest at these points. The gaps are to be considered as filled with the name of the color next adjoining; for example, phenolphthalein is colorless for all pH concentrations greater than 10^{-9}, and red in all lower than this.

From the foregoing discussion and from a consideration of the table of indicators in Fig. 25, it is clear that hydrogen-ion concentrations may be determined through a certain range by means of indicators. One selects an indicator that is believed to change color at a pH near that of the unknown solution, since this is the type of indicator that is most likely to give a satisfactory result. The unknown solution is treated with a few drops of the selected indicator, and the color obtained is compared with that produced with the same amount of the same indicator in a solution of known hydrogen-ion concentration. (It is evident that a series of these solutions of known pH must be kept on hand.) If the tint produced in the unknown solution is identical with that of a solution of known pH, then the acidities of the two are the same. The choice of indicators for the purpose of determining the hydrogen-ion concentration is somewhat different from that for titration purposes; i.e., in this case we need an indicator that shows a very gradual change in color through a given range, and one that is not excessively sensitive to neutral salts and proteins. Clark and Lubs[10] recommend the following carefully selected series of indicators as being especially adapted to the colorimetric method in connection with their series of standard buffer mixtures (see below).

Preparation of Indicator Solutions.—The indicators recommended by Clark and Lubs may be obtained both in the form of a dry powder and in stock solutions. In cases where the acidity of the free acid dye in the indicator solution does not interfere with accuracy and when alcohol is not objectionable, Clark and Lubs claim that the alcoholic

[10] J. Bact., **2**, 1, 1917.

Indicator	Colour of Indicator with Hydrogen ion Concentration of N X.														Concentration of Indicator Solution	Drops added to 10 c.c. of test
	2	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}	10^{-11}	10^{-12}		
Methyl-violet 6B. = Crystal-violet	Gold Yellow	Green	Green Blue	Blue	Violet										0.05%	3·8
Tropeolin oo = Diphenylamine-orange			Red	Flesh	Yellow										0.01%	3·5
Benzyl-anilino-azo-benzene. Sorensen p. 99				Red	Orange										0.02% in 50% alcohol	5·10
(Dimethyl-amino-azo-benzene)				Red	Flesh	Gold Yellow									0.01% in 80% alcohol	5·10
Methyl-orange = Tropeolin D.				Red	Orange Red	Orange	Yellow								0.01%	3·5
Congo Red					Blue	Violet	Scarlet								0.01%	3·5
Methyl Red					Violet Red	Red	Orange	Yellow							0.02% in 60% alcohol	4
Paranitrophenol							None	Faint Green	Green Yellow						0.04% in 6% alcohol	3·20
Neutral Red	Blue	Blue Violet	Red						Rose	Orange Yellow					0.01% in 50% alcohol	10·20
α-Naphthol-phthalein									None	Greenish	Blue				0.04% in 60% alcohol	4·12
Tropeolin ooo I = α-naphthol-orange									Yellow	Orange Red	Red				0.01%	4·10
Phenol-phthalein										None	Rose	Red			0.05% in 50% alcohol	3·20
Thymol-phthalein												None	Blue		0.04% in 50% alcohol	3·20
Tropeolin o = Resorcinol-yellow													Green Yellow	Orange	0.01%	5·10

FIG. 25.—Table of Indicators, Compiled from those of Salm, Sörensen and Bayliss.

TABLE XVIII

CLARK AND LUBS SELECTED INDICATORS

Indicator	Concentration, per cent	Color Range	pH Range
Thymol sulphon phthalein (thymol blue) acid range.................... .	0.04	Red-Yellow	1.2-2.8
Tetra bromsulphon phthalein (brom phenol blue).....	0.04	Yellow-Blue	3 0-4.6
Ortho carboxy benzene azo dimethyl aniline (methyl red)......	0.02	Red Yellow	4.4-6.0
Di bromo ortho cresol sulphon phthalein (Brom cresol purple)............... ..	0.04	Yellow-Purple	5.2-6.8
Di bromo thymol sulphon phthalein (brom thymol blue).........	0 04	Yellow-Blue	6.0-7.6
Phenol sulphon phthalein (phenol red) ...	0.02	Yellow-Red	6.8-8.4
Ortho cresol sulphon phthalein (cresol red).	0.02	Yellow-Red	7.2-8.8
Thymol blue—alkaline range...........	0.04	Yellow-Blue	8.0-9 6
Ortho cresol sulphon phthalein......... .	0 02	Colorless-Red	8.2-9.8

solutions may be used; but for most work they prefer to use aqueous solutions of the alkali salts in concentrations which may be conveniently kept as stock solutions. To prepare these stock solutions they recommend rubbing 0.1-gram portions of the dry powder in an agate mortar with the following quantities of N/20 NaOH: phenol red, 5.7 c.c.; brom-phenol blue, 3.0 c.c.; cresol red or meta-cresol purple, 5.3 c.c.; thymol blue, 4.3 c.c.; brom-thymol blue, 3.2 c.c.; methyl red, 7.4 c.c. When the indicators are completely dissolved, the volume of each is made up to 25 c.c. with water. These are the stock solutions. When ready for use, those indicators that are supposed to be 0.04 per cent should be diluted 10 times. (A small portion is diluted this amount.) Those indicators that are supposed to be 0.02 per cent should be diluted 20 times.

Standard Solutions of Definite pH.—In order that the indicator method may be satisfactorily applied in the determination of hydrogen-ion concentrations, one must have on hand a series of standard solutions the pH values of which are definitely known. These can be produced by using buffer mixtures whose pH is not readily altered (see p. 196). It will be remembered that buffer solutions are such as contain only a small concentration of free H^+ ions and OH^-, but a large reserve

of potential ions. Consequently, when more acid or alkali is added, combination with the potential ions takes place, but the pH of the solution undergoes but little change. Buffer mixtures, therefore, are indispensable in the preparation of standard solutions of definite pH. In short, solutions of known hydrogen-ion concentration can be prepared by first preparing a suitable set of stock buffer solutions, and then mixing these in certain proportions. In this way, solutions of different hydrogen-ion concentrations can easily be prepared. Thus, a mixture of 7 volumes of 0.20 N sodium acetate and 3 volumes of 0.20 N acetic acid solution gives a mixture with a hydrogen-ion concentration of $1 \times 10^{-5} (pH\ 5.0)$.

A number of such buffer solutions have been suggested and used. For instance, Clark and Lubs have prepared a series using the following solutions: 0.2 M acid potassium phthalate, 0.2 M acid potassium phosphate, 0.2 M boric acid in 0.2 M potassium chloride, 0.2 N hydrochloric acid, and 0.2 N sodium hydroxide. A definite volume (50 c.c.) of the standard solutions—acid potassium phthalate, acid potassium phosphate, or boric acid in potassium chloride—is measured into a 200-c.c. measuring flask; then a given amount ($\times$ c.c.) of standard NaOH or HCl is added, and the volume brought up to the 200-c.c. mark with distilled water. The pH values obtained with different mixtures are given below. In the preparation of the original solutions of standard pH, the exact hydrogen-ion concentration was determined electrometrically. (The pH values for the mixtures given below were obtained in this way.) For ordinary work, however, it is sufficient to use the concentrations given and to accept the pH values as correct. For more exact work, one should check one's standard solutions electrometrically.

Method of Determining the pH Value.—With a pipette, transfer 2 c.c. of the unknown solution to a 20-c.c. measuring flask and add water to make up to a total volume of 20 c.c. Mix and add the specified amount of indicator, using phenol red first. Compare with the set of standards for this indicator. In case the indicator is beyond the limits of phenol red on either side, repeat with the next indicator in order, until the unknown has been correctly matched against one of the standards.

7. Hydrogen-ion Content of the Blood.—The pH of blood may be determined either by the hydrogen-electrode method or by the indicator method. The former method is the more accurate, but it

TABLE XIX

pH OF STANDARD BUFFER MIXTURES

Series A.—To 50 c.c. of 0.2 M acid potassium phthalate add the indicated number of c.c. of 0.2 N HCl and dilute to 200 c.c.   Indicators recommended: Thymol blue or brom-phenol blue.   Use 5 drops.

HCl	pH	HCl	pH	HCl	pH	HCl	pH
c.c.		c.c.		c.c.		c.c.	
46.7	2.2	32.9	2.6	20.3	3.0	9.9	3.4
42.5	2.3	29 6	2.7	17.7	3.1	7.5	3.5
39 6	2.4	26.4	2.8	14.7	3.2	6.0	3.6
37.0	2.5	22.8	2.9	11.8	3.3	4.3	3.7
						2.6	3.8
						1.0	3.9

Series B.—To 50 c.c. of 0.2 M acid potassium phthalate, add the indicated number of c.c. of 0.2 N NaOH, and dilute to 200 c.c.   Indicators recommended: Brom-phenol blue or brom-cresol purple.   Use 3 drops.

NaOH	pH	NaOH	pH	NaOH	pH	NaOH	pH
c.c.		c.c.		c.c.		c.c.	
0.4	4.0	12.2	4.6	30.0	5.2	43.0	5.8
2.2	4.1	14.6	4.7	32.5	5.3	44.6	5.9
3.7	4.2	17.7	4.8	35.5	5.4	45.5	6.0
5.2	4.3	21.0	4.9	37.7	5.5	46.2	6.1
7.5	4.4	23.9	5.0	39.9	5.6	47.0	6.2
9.6	4.5	27.2	5.1	41.9	5.7	48.1	6.3

Series C.—To 50 c.c. of 0.2 M acid potassium phosphate, add the indicated number of c.c. of 0.2 N NaOH, and dilute to 200 c.c.   Indicators recommended: Phenol red or brom-cresol purple.   Use 3 drops.

NaOH	pH	NaOH	pH	NaOH	pH	NaOH	pH
c.c.		c.c.		c.c.		c.c.	
3.7	5.8	12.6	6.4	29.6	7.0	42.8	7.6
4.7	5.9	16.0	6.5	32.5	7.1	44.2	7.7
5.7	6.0	17.8	6.6	35.0	7.2	45.2	7.8
7.4	6.1	21.0	6.7	37.4	7.3	46.0	7.9
8.6	6.2	23.7	6.8	39.5	7.4	46.8	8.0
10.2	6.3	26.5	6.9	41.2	7.5		

TABLE XIX—*Continued*

Series D.—To 50 c.c. of 0.2 M boric acid in 0.2 M potassium chloride, add the number of c.c. of 0.2 N NaOH indicated below, and dilute to 200 c.c. Indicators recommended: Thymol blue or cresol red. Use 3 drops.

NaOH	pH	NaOH	pH	NaOH	pH	NaOH	pH
c.c.		c.c.		c.c.		c.c.	
2.6	7.8	8.5	8.4	21.3	9.0	36.9	9.6
3.3	7.9	10.4	8.5	24.3	9.1	39.0	9.7
4.0	8.0	12.0	8.6	26.7	9.2	40.8	9.8
4.8	8.1	14.3	8.7	30.0	9.3	42.5	9.9
5.9	8.2	16.3	8.8	32.0	9.4	43.9	10.0
7.3	8.3	19.0	8.9	34.5	9.5		

requires considerable experience. In most cases the indicator method is quite satisfactory, if a few simple precautions are observed; for example, it is necessary to dilute blood-plasma about ten to twenty times, or else to dialyze it through a collodion sac in order that the blood-proteins may be reduced in quantity. Dilution does not greatly alter the hydrogen-ion concentration, and dialysis likewise makes but little change, provided exposure to air does not occur. A number of corrections, however, have to be applied in diluting blood-plasma. In the first place it is necessary to dilute the blood-plasma with a neutral salt solution, such as 0.9 per cent NaCl, and this introduces a "salt error," which has to be corrected. In the second place, there is a "temperature error" for the indicator, since the color of an indicator depends somewhat on the temperature. This latter error can be avoided by making all measurements at the same temperature.

Cullen[11] has standardized a method using phenol red for the indicator. He dilutes the blood-plasma twenty times with 0.9 per cent NaCl solution, so that the salt error is comparable to that of the original plasma. He also employs tubes of standard color in phosphate mixtures, and subtracts 0.225 from the reading to get the true pH.

Van Slyke,[12] while studying acidosis, made a very careful study of the pH of blood, and found that the maximum normal range of variation of reaction in different individuals is indicated by pH 7.30–7.50. Under extreme abnormal conditions the pH may fall slightly below 7.0, or may rise as high as 7.8. With the former condition *coma* occurs, while with the latter *tetany* is generally present.

[11] J. Biol. Chem., **52**, 501–515, 1922.   [12] J. Biol. Chem., **48**, 153–176, 1921.

CHAPTER XI

THE COLLOIDAL STATE OF MATTER

GENERAL INTRODUCTION

1. Crystalloids and Colloids.—In an article published in *Philosophical Transactions* in 1861, an English chemist, Thomas Graham, differentiated all instances into two classes; those that rapidly diffuse through organic membranes (vegetable parchment paper, fish bladders, etc.), and those that either do not diffuse through such membranes or diffuse very slowly through them. The substances belonging to the first class readily form crystals from saturated solutions. They include such compounds as salt, sugar, urea, etc. Those belonging to the second class are not generally found in the crystalline condition. They include such compounds as silicic acid, soluble alumina, and certain organic compounds, viz., gum arabic, caramel, starch, dextrin, tannin, albumin, etc. Graham called the former class *crystalloids*, and the latter *colloids* (from *colla*, "glue").

2. Graham's View on Colloids.—Since the time of Graham it has been found that most substances can be obtained in either the crystalloidal or the colloidal state, the state of aggregation depending on the manner of manufacture. Furthermore, both states of aggregation are often found in the same solution. The two states can be separated by diffusion; in fact, Graham used this method for separating the one class from the other, and he termed this separation *dialysis*, and the apparatus used in effecting such separations a *dialyzer*. When a mixture of the two states of aggregation is placed in a dialyzer, the crystalloids pass through and the colloids remain behind. The results, however, are not quantitative.

A fair idea of Graham's views on colloids may be obtained from the following quotations from his writings:

"I may be allowed to advert again to the radical distinction assumed in this paper to exist between colloids and crystalloids in their

intimate constitution. Every physical and chemical property is characteristically modified in each class. They appear like different worlds of matter, and give occasion to a corresponding division of chemical science. The distinction between these kinds of matter is that subsisting between the material of a mineral and the material of an organized mass.

Again: "The phenomena of the solution of a salt or crystalloid probably appear in the solution of a colloid, but greatly reduced in degree. The process becomes slow; time, indeed, appearing essential to the colloidal changes. The change of temperature, usually occurring in the act of solution, becomes barely perceptible. . . . The colloid, although often dissolved in large proportion by its solvent, is held in solution by a singularly feeble force. Hence colloids are generally displaced and precipitated by the addition to their solution of any substance from the other class. Of all the properties of liquid colloids, their slow diffusion in water and their arrest by colloidal septa are the most serviceable in distinguishing them from crystalloids. Colloids have a feeble chemical reaction. . . . It is difficult to avoid associating the inertness of colloids with their high equivalents, particularly where the high number appears to be attained by the repetition of a smaller number. The inquiry suggests itself whether the colloid molecule may not be constituted by the grouping together of a number of smaller crystalloid molecules, and whether the basis of colloidality may not be this composite character of the molecule."

This paper concludes with a discussion of "Osmose," and the views expounded are essentially those that we hold to-day:

"It now appears to me that the water movement in osmose is an affair of hydration and of dehydration in the substance of the membrane or other colloid septum. The surface of the membrane, being in contact with pure water, tends to hydrate itself to a higher degree than the inner surface does, the latter surface being supposed to be in contact with a saline solution." Where the membrane comes in contact with the solution, "the degree of hydration is lowered, and the water must be given up by the inner layer of the membrane and it forms the osmose."

Our present-day views on the passage of water through semi-permeable membranes in osmotic-pressure procedures are essentially the same as the views held by Graham.

Graham saw the importance of colloids, not only for inanimate

matter, but for all living matter. This can best be seen by quoting his own words:

"The colloidal is, in fact, a dynamical state of matter; the crystalloidal being the statical condition. The colloid possesses *energeia*. It may be looked upon as the probable primary source of the force appearing in the phenomena of vitality. To the gradual manner in which colloidal changes take place (for they always demand time as an element), may the characteristic protraction of chemical-organic changes also be referred."

This is prophetic of the present-day conception of the importance of the colloidal state of matter in life processes, for we now know that all the phenomena of life are connected with processes that take place only in colloid systems.

3. Colloids a State of Aggregation and not a Distinct Class of Substances.—The most important advance in our knowledge of colloids that has been made since Graham's fundamental work has been the proof, now generally considered conclusive, that there are no "colloids" in the sense that Graham used the term. In other words, colloids are not a distinct class of substances endowed with peculiar properties, for, in all probability, all substances may be made to assume the colloidal state. Certain substances tend to form a colloidal solution with one solvent, and dissolve as crystalloids in another. For example, soap dissolves in alcohol as a crystalloid, i.e., a soap forms a true solution in alcohol, but it dissolves in water in the form of a colloidal suspension. Common salt and many other inorganic substances, usually obtained in the crystalline form, may be also obtained in colloidal solution if they are produced by chemical reactions in the midst of liquids in which they happen to be insoluble.

Consequently, it is not strictly correct to speak of a substance as being a *colloid*; it should be designated as being in the *colloidal state.*

A further point, which has been established in late years, is that colloidal solutions are not solutions in the general use of the term. A true solution may be considered as a *homogeneous* mixture, and therefore consists of a single phase. A colloidal solution, on the other hand, is strictly *heterogeneous*; that is, it consists of two phases at least. In a true solution the dissolved substance is considered to be present either in the molecular or the ionic condition. This fact may be shown by the determination of its molecular weight by the freezing-point or the boiling-point method. If the solution contains particles that are larger

than single molecules or ions, we have a type of solution that is either a colloidal solution or a mechanical suspension. The phase that is divided into small separate particles, either liquid or solid, is known as the *disperse phase*; and the liquid in which it is distributed is the *dispersion medium*.

4. Sols and Gels.—Graham called the colloidal solutions obtained with water as the solvent "hydrosols," or simply "sols," and those that formed a gelatinous mass in water "hydrogels," or "gels." When other solvents, such as alcohol or glycerol, were used, the colloidal solutions produced were termed "alcosols," "alcogels," "glycerosols," and "glycerogels," respectively.

PREPARATION OF COLLOIDAL SOLUTIONS

5. Methods of Preparing Colloidal Solutions.—A fairly large number of methods have been devised for preparing colloidal solutions with compounds that ordinarily do not appear in colloidal form. These methods admit of easy classification, in that they fall into two general classes: chemical methods and electrical methods.

The chemical methods include reduction methods, oxidation methods, hydrolytic methods, double-decomposition methods, methods involving replacement of solvent, "protective methods," and "peptization" methods.

The electrical method has come prominently into play in recent years.

Gelatin, gum, some proteins, and certain other substances form colloidal solutions on simple solution in water. Other substances form colloidal solutions when dissolved in dilute saline solutions.

6. Preparation of Colloidal Solutions by Chemical Reduction.— The reduction reactions have to do primarily with the reduction of salts of the metals by certain reducing agents, such as hydrogen gas, carbon dioxide, phosphorus, tannin, etc.

The reduction methods were the first to be used in preparing colloidal solutions. Thus, at the very beginning of the nineteenth century, gold sols were prepared by reducing salts of gold with mild reducing agents. The product was not called colloidal gold, however.

Faraday, in 1857, prepared a colloidal gold solution having a red color by treating gold chloride with an ethereal solution of phosphorus. He expressed the opinion that the gold was suspended in the liquid in

an extremely fine state of division. Similarly, by the action of certain reducing agents on soluble silver salts, Carey Lea (1887) obtained colored solutions containing colloidal silver.

The history of the preparation of purple of Cassius is interesting. If a mixture of stannic and stannous chlorides is added to a very dilute solution of gold chloride, hydrated stannic oxide is precipitated and the gold chloride is reduced to a metal. The precipitate of stannic hydroxide may have tints varying from red to violet according to the concentration and composition of the solution. This precipitate is called purple of Cassius because A. Cassius wrote a pamphlet, entitled " De Auro," describing its preparation, in 1685. The mode of making this substance, however, was known to B. Valentine in 1603. The nature of this purple of Cassius was the subject of much speculation in former years, and it is only since colloidal chemistry has come to the front that it has been understood. In short, the purple of Cassius is now known to be composed of colloidal gold and colloidal stannic acid, and these two components remain together throughout all the reactions unless the purple of Cassius is actually destroyed.

Stable sols of platinum have also been prepared by reducing its salts with acrolein.

A few colloidal solutions have also been made by the opposite of reduction methods, viz., oxidation methods; but these are not sufficiently important for any detailed discussion here.

7. Preparation of Colloidal Solutions by Hydrolysis.—The hydrolysis of salts may be considered as being a special case of double decomposition in which water is one of the substances evolved. The reaction may be represented by the following equation:

$$Ct' - An^\cdot + H'OH^\cdot \rightleftarrows H' - An^\cdot + Ct' - OH^\cdot,$$
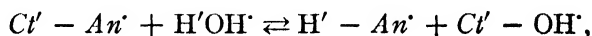
where Ct' and $An^\cdot$ stand for cation and anion, respectively, of the dissolved substance. When, in the course of hydrolysis, one of the products is practically insoluble, conditions are favorable for obtaining a stable colloidal solution. For example, if 1 c.c. of a filtered 33 per cent solution of ferric chloride is added to 100 c.c. of boiling distilled water, a reddish-brown solution of colloidal ferric hydroxide is obtained.

8. Preparation of Colloidal Solutions by Double Decomposition.— Under this heading we may include most of the chemical condensation methods, which are not included under reductions, oxidations, or

hydrolyses. One of the double-decomposition products must be nearly insoluble. We can take the preparation of colloidal cadmium sulphide as an illustration of this method. If we treat a very dilute solution of cadmium chloride with hydrogen sulphide, the solution gradually becomes milky and finally develops a yellowish color with reddish surface. The excess of hydrogen sulphide must be removed by a current of nitrogen or by boiling.

9. Preparation of Colloidal Solutions by Protective Media.— Many substances, which if formed in pure water would tend to precipitate out, will be retained in a finely divided state of suspension provided some viscous medium is present. Small quantities of gelatin, gum arabic, dextrin, starch, albumin, etc., seem to exert some sort of protective action on the small particles as they are formed, thereby preventing their clumping into larger masses. Furthermore, colloidal solutions of certain inorganic salts may be prepared by dissolving these salts in such fluids as glycerol and then pouring them into water.

10. Preparation of Colloidal Solutions by Mechanical Dispersion. — Certain substances, such as the natural gums, which are insoluble in water, may be converted into hydrosols (colloidal solutions in water) by dissolving them in some other liquid and then pouring into a large volume of water. For example, a colloidal solution of gum mastic may be prepared by dissolving a small quantity of this substance in ethyl alcohol and then adding it to a large volume of water. A milky-white permanent colloidal solution of gum mastic is obtained. It is, however, very sensitive to electrolytes.

11. Preparation of Colloidal Solutions by " Peptization."—Under the influence of certain foreign substances (sometimes traces suffice), hydrogels or precipitates may be changed into hydrosols. The term "peptization" was introduced by Graham to express the transformation of a gel into a sol. At the present time, we understand a peptizer to be a substance that, if sufficiently concentrated, is capable of effecting the solution of a solid that is insoluble in its dispersion medium. A typical example of peptization is afforded by silver chloride, which forms a sol on prolonged digestion with a solution containing either Ag· or Cl′. It is apparent that the rate of peptization can be controlled by dilution of the peptizer, and that when the sol stage is reached it is possible to remove most of the peptizer from the mixture by dialysis.

Insoluble precipitates are often converted into hydrosols by the

removal of an excess of electrolytes by prolonged washing or dialysis. A familiar example of this type of peptization is furnished by the tendency of many precipitates to run through the filter after too prolonged washing with water.

12. Preparation of Metal Sols by Electrical Dispersion Methods.— It was known, even at the time of Sir Humphry Davy, that when a heavy electrical discharge is passed between metal poles the metal particles are torn off in a fine state of division.

Utilizing this fact, Bredig[1] devised a method for preparing colloidal solutions of many of the most resistance metals. He found that, when an (continuous-current) arc was produced under water between electrodes of the noble metals, the liquid became deeply colored and showed all the properties of colloidal solutions known at that time. Some metals yielded only coarse dispersions, while others were oxidized, hydroxide sols being formed. As a general rule, a small quantity of alkali favored the production of stable sols.

Bredig's method worked very well in water; but when other liquids, especially the organic solvents, were employed, it did not give satisfactory results. This was due to the fact that the liquids themselves were prone to undergo decomposition, liberating carbon which contaminated the sol in question.

The dispersion by the electric arc was further developed by Svedberg.[2] He overcame some of the difficulties encountered by Bredig and succeeded in preparing sols in other solvents than water. Instead of the continuous current he used various other types of discharge, e.g., oscillating; he employed electrodes of metal that do not disintegrate, such as iron or aluminium, between which the metal to be dispersed was suspended as foil or fragments; and he carried out the dispersion in a number of organic liquids.

He prepared sols of all the alkali metals in ethyl ether, as well as a large number of metal sols in isobutyl alcohol. He also obtained sols of selenium, carbon, silicon, and of a number of minerals.

GENERAL CONSTITUTION OF COLLOID SYSTEMS

13. Heterogeneity of the Colloidal State of Matter.—Colloidal solutions are often as clear and transparent to the naked eye as true solutions; but they may be recognized by certain well-established tests.

[1] Zeit. phys. Chem., **31**, 258, 1899.
[2] Ber. d. chem. Gesell., **38**, 3616, 1905; **39**, 1705, 1906; Koll. Zeit., **1**, 229, 257, 1907.

(*a*) *By Dialysis.*—One of the simplest and most convenient methods of ascertaining whether or not a colloidal dispersoid is present in a solution is that of dialysis as used by Graham. As a dialyzer he used a piece of parchment paper fastened between two hoops, forming a sort of tray which could be immersed in water or other liquids. The solution to be tested for its colloidality is placed inside the dialyzer, which is then immersed in water or other liquid. If the solution on the inside of the dialyzer passes through the walls of the dialyzer to the liquid on the outside, the solution is probably not colloidal in nature.

Collodion thimbles also make excellent dialyzers. These are made by coating the surfaces of test-tubes, Erlenmeyer flasks, etc., with a solution of collodion, then immersing in water and removing the tough membrane from the glass vessel.

Most crystalloids will pass through such a container, but colloids will not. For example, if we place a mixture of glucose solution and starch solution in one of the above dialyzers, the glucose will dialyze through but the starch will not.

(*b*) *By Turbidity.*—If some colloidal solution is placed in a perfectly clean, thin-walled test-tube, and held against an opaquely black background, e.g., black velvet, the liquid may assume a grayish-white appearance.

(*c*) *Tyndall Phenomenon.*—If a bright beam of light is passed through a colloidal solution contained in a vessel with parallel sides, and the vessel is then viewed from the side, it will appear turbid in the path of the light. If a converging lens is used, the turbidity will be cone-shaped. A "control" should be made with dust-free distilled water.

It should be noted that it is not the presence of a number of more or less evident particles, which might be recognized either macroscopically or microscopically, that distinguishes a colloidal solution from a true one. It is rather the *intensity* of the unbroken light-cone passing through the solution which is the determining factor. At any rate, it is safe to say that liquids that do not exhibit the Tyndall light-cone, or show it only in high concentrations, are molecular-disperse solutions, since it is extremely rare that a colloidal solution does not give a positive Tyndall effect.

(*d*) *The Ultramicroscopic Test for Colloidal Solutions.*—The ultra-microscope, like the Tyndall phenomenon, is a development of Faraday's observations on colloidal gold. Faraday found that when

a bright beam of light is sent through certain solutions of gold and then is looked at from the side, the light beam appears as a yellow track. He concluded that this is due to light diffracted from solid particles of gold, even though these particles could not be seen by the naked eye or even by using a powerful microscope. Tyndall later made the observation that the light was often polarized, which means that in such cases the particles have dimensions similar to the wave-length of the illuminating beam. Siedentopf and Zsigmondy,[3] at a
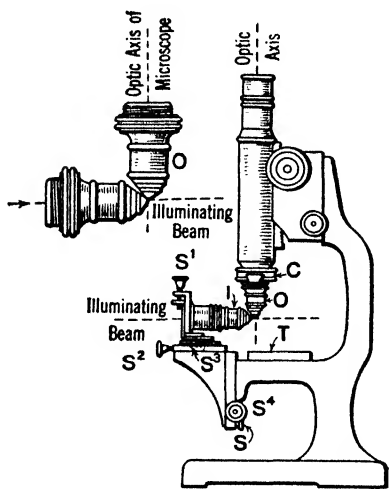
FIG. 26.—The Ultramicroscope.

still later date, arranged a micro-scope to observe the beam, and by this means were able to see the separate particles in the solution by means of their diffraction images (see Fig. 26). It occurred to them that if the colloidal solutions were greatly diluted and the beam examined by a microscope, placed perpendicularly to its track, so as not to receive the direct light, the diffraction images of the separate particles would be visible. They consequently constructed an instrument for this purpose, which they called an "ultramicroscope." A powerful beam of light, issuing from a horizontal slit, is brought to a focus within the colloidal solution under examination, by means of a microscope objective, and this image is viewed through a second microscope, the axis of which is at right angles to the path of the beam. If the liquid is optically void, the field of view in the microscope will appear dark; but if colloidal particles are present in the liquid, minute points of light will be observed against the dark background. these points of light will be observed to be in more or less rapid motion; in fact, colloidal solutions usually appear to be swarming with brilliantly colored particles (p. 246). It must be emphasized that it is not the particles themselves that are visible, but the light that is reflected from them.

Particles visible in an ordinary microscope are usually described as

[3] Colloids and the Ultramicroscope, by R. Zsigmondy. Translated by Alexander. John Wiley & Sons.

microns; those that can be made visible by the ultramicroscope, as submicrons; and those that are too small to be rendered visible even by the latter, as amicrons. These differences, however, are not determined by size alone.

(e) *Size of the Particles.*—The average size of the particles in a colloidal solution can be estimated indirectly by counting the number in a given volume and determining the total amount of substance by analysis.

Zsigmondy has suggested the following formula for the determination of the size of the particles:

$$L = \sqrt[3]{\frac{A}{S \cdot N}}.$$

L is the length of the side of one particle, *A* is the mass of the dispersed substance in a given unit volume, *N* is the number of submicrons in the same volume, and *S* is the specific gravity of the dispersed substance. It should be noted that this formula only holds under the assumptions that the particles are cubical in shape, that they are uniformly distributed, and that they are visible and exactly alike. Zsigmondy claims that by closely following directions one can ultimately obtain a close approximation to the average size of the particles. There are other methods of obtaining the same results.

It has been shown that the particles vary greatly in magnitude, depending on the nature and mode of preparation of the colloidal solution. They range from particles that can be seen with an ordinary microscope to those that are invisible even in the ultramicroscope. Particles that are visible in an ordinary microscope have diameters exceeding 250 $\mu\mu$, and those that are invisible in the ultramicroscope have diameters less than 6 $\mu\mu$. (A micron, represented by μ, is equal to one-thousandth of a millimeter (0.001 mm.); and a millimicron, represented by $\mu\mu$, is equal to one-thousandth of a micron, or to one-millionth of a millimeter.)

Some idea of the size of a typical colloidal particle, as compared with a typical substance that is easily visible in an ordinary microscope, will be obtained from the statement that the diameter of human red blood corpuscles is about 8 μ, and that most micrococci have diameters of about 1 μ, while the particles of a colloidal gold sol have diameters of 7–15 $\mu\mu$. The diameter of a molecule of hydrogen has been calculated to be about 0.1 $\mu\mu$. An idea of the relative sizes of

colloidal particles and molecules in comparison with a blood corpuscle or a bacterium is given in Fig. 27.

It should be evident from the foregoing data that the peculiar and distinctive properties of colloidal solutions are due to the enormous development of surface. The difference between the properties of a
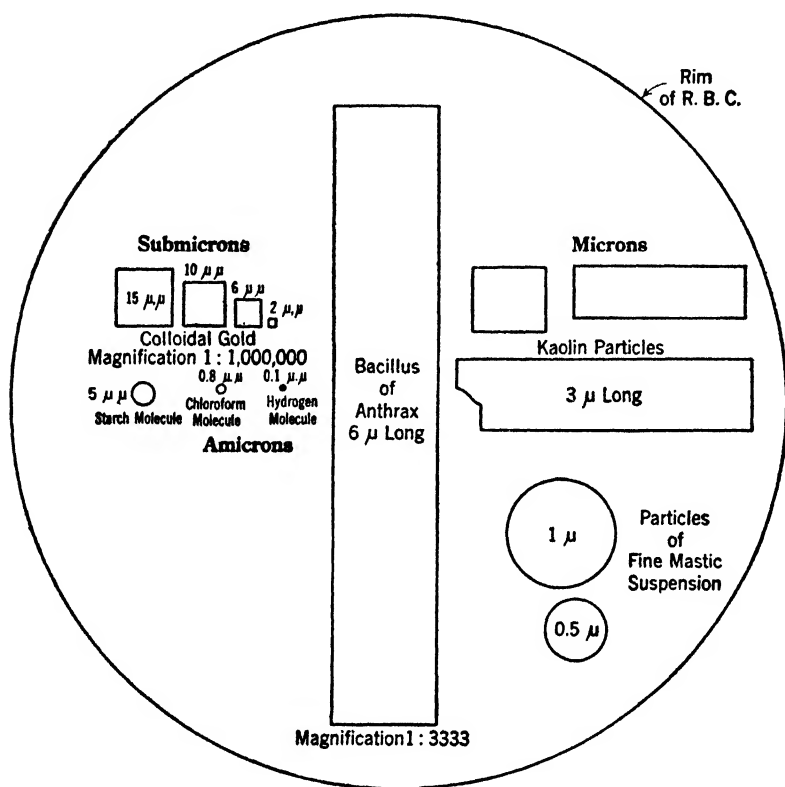
Fig. 27.—Diagram from W. Ostwald, showing the Relative Size of Various Molecules, Colloidal Dispersoids, etc. Compared with a Red Blood Cell, and an Anthrax Bacillus.

solution of colloidal platinum, for example, and a bullet of platinum in water, depends on the fact that the surface of the platinum in the former case is many million times that of the same weight of platinum in the bullet form. Table XX makes clear the enormity of the increase in effective surface that takes place when a sphere is divided into smaller particles.

TABLE XX

Length of Radius	Number of Spheres	Total Surface
1 cm. Small bullet.............	1	12.6 sq. cm.
0.1 mm. ⎫	10^6	1260.0 sq. cm.
0.01 mm. ⎬ Coarse suspensions ⎰	10^9	1.26 sq. meters
1 μ ⎭	10^{12}	12.60 sq. meters
0.1 μ ⎫	10^{16}	126.00 sq. meters
0.01 μ ⎬ Typical colloids... ⎰	10^{18}	1,260.00 sq. meters
1 μμ ⎭	10^{21}	12,600.00 sq. meters
0.1 μμ True solution.............	10^{24}	126,000.00 sq. meters

These figures demonstrate the extraordinary amount of adsorbing surface of a small amount of matter highly dispersed. It is also evident that properties dependent on osmotic pressure, diffusibility, etc., are exhibited by all dispersoids, whether ions, molecules, colloidal particles, coarse suspensions, or pieces of matter; but some of these properties are much more pronounced when the dispersoids are of large dimensions, and others when they are small. For example, the phenomena due to surface development become suppressed when the particles attain the dimensions of molecules and ions, since this property is only appreciable when the dispersoids have the condition of matter in mass. Not only do colloidal solutions exhibit the mass phenomenon, but the suspended particles possess a total surface that is enormous in relation to the total mass present.

14. Classification of Colloids.—It has been found necessary to "fix" the limits of the size of particles that will be accepted as belonging to the colloidal state. Since the limit of microscope vision is about one-ten-thousandth of a millimeter, or 0.1 μ, this value is arbitrarily taken as the dividing line between coarse suspensions and colloidal dispersoids. Now, true molecular dispersoids have diameters of one ten-thousandth to one one-millionth of a millimeter (0.1 μ to 1 μμ). Hence, one one-millionth of a millimeter is accepted as the dividing line between colloidal and molecular dispersoids. Note the following diagram:

DISPERSED SYSTEMS (FROM OSTWALD)

Coarse Dispersoids	Colloids	Molecular Dispersoids
Particles greater than 0.1 μ; do not pass through filter papers; microscopically analyzable.	Increase, in degree of dispersion, 0.1 μ to 1 $\mu\mu$	Particles smaller than 1 $\mu\mu$; pass through filter paper; cannot be analyzed microscopically; do diffuse and dialyze.
	Pass through filter papers; cannot be analyzed microscopically; do not diffuse or dialyze.	

The three divisions are further divided as follows:

1. Coarse dispersoids; suspensions and emulsions.

2. Colloids; suspensoids and emulsoids.

3. Molecular dispersoids; crystalloidal solutions, both electrolytes and non-electrolytes.

Suspensions and suspensoids are systems in which solid, non-deformable particles are suspended in a liquid.

Emulsions and emulsoids are systems of two liquids insoluble in each other; they consist of liquid, deformable particles of one liquid suspended in another liquid with which they do not mix.

Most of the natural colloidal solutions are emulsoids, and most of the artificial colloidal solutions are suspensoids.

From a biological standpoint the emulsoids are much more important than the suspensoids. Emulsoids include suspensions in water of proteins, starch, glycogen, agar, soaps, etc. The two types of colloidal suspensions possess certain properties in common, but differ widely in other characteristics; it will be necessary, therefore, to consider some of their properties separately.

MECHANICAL PROPERTIES OF DISPERSOIDS

15. Ultrafiltration.—It will be noted in the above classification that coarse suspensions are characterized as being retained by filter papers, whereas colloidal suspensoids, emulsoids, and molecular dispersoids are said to pass through filter papers. There is no sharp dividing line, however; the rate of filtration varies widely in all the types mentioned. Perhaps the greatest variation is among the emulsoid sols, owing to the wide range of viscosity among them.

Bechhold[4] has demonstrated that it is possible to form, on filter paper or wire gauze, collodion films with varying amounts of acetic acid, or gelatin films hardened with formaldehyde, such that pores of almost any desired size can be obtained. By this means it is not only possible to separate suspensoids from their dispersion media, but also to effect the concentration of emulsoids. Ultrafiltration, however, is a very complex process, since it involves the phenomena of adsorption and dialysis, as well as the ordinary process of mechanical separation. As we shall learn later, colloidal particles are electrically charged at their surfaces. Some are electropositive, some electronegative; and in the case of colloidal ampholytes, they are either electropositive or electronegative, depending on the reaction of the media. Now, if we attempt to filter an electropositive sol, such as ferric hydroxide sol, we encounter the following difficulty: The filter becomes negatively charged in contact with water; consequently, on the entrance of the positively charged sol into the pores of the filter, the colloid is immediately discharged and the disperse phase precipitated. After the pores of the filter become partially stopped with particles of the colloid, the sol will pass through unchanged.

Ultrafilters are also used to produce optically void water for work with the ultramicroscope, and to produce absolutely sterile water for bacteriological purposes. The behavior of mixtures is also very interesting. If we have two substances that do not tend to adsorb each other, provided the particles are of different sizes, a separation can be effected. Furthermore, by the addition of a protective colloid to certain suspensoids, a further dispersion takes place and the suspensoid will pass through a filter which retained it before.

16. Diffusion and Dialysis.—Diffusion and dialysis experiments provide us with two of the simplest methods for distinguishing between molecular-disperse (true) and colloidal solutions. These might be called the "classical" methods for the qualitative analysis of these solutions, for it was by them that Graham, in 1861, first distinguished between the "states" of different solutions and thus introduced the concept "colloid." If we recall the relations already pointed out when we were considering true solutions, between diffusion and osmotic pressure, viz., that all diffusion is caused by osmotic pressure, we should expect to find all colloids diffusing very slowly (owing to the size of the particles), if diffusing at all. Such was shown to be the

[4] Zeit. phys. Chem., **60**, 257, 1907; **64**, 328, 1908.

case by Graham in his classical researches on colloids. The relative rates of diffusion of a number of colloids, as compared with certain organic crystalloids, are given below:

	Diffusing Constant
Urea	1 01
Glucose	0.57
Pepsin	0.063
Albumin	0 054
Emulsion	0.036

These figures show the relative slowness with which colloids diffuse; and these are emulsoids, which, as we shall learn, have greater osmotic pressures than suspensoids. In other words, since suspensoids have smaller osmotic pressures than emulsoids, we should expect them to diffuse even more slowly, and here again experiment proves the prediction. In fact, there are times when it would seem to be doubtful whether certain suspensoids, at least under certain conditions, diffuse at all.

Many simple experiments based on diffusion have been devised for the purpose of distinguishing between true and colloidal solutions. Ostwald uses 5 per cent gelatin solutions or 2 per cent agar-agar solutions, which while hot are poured into test-tubes until these are half-filled, and then allowed to congeal, care being taken that the gelatin and agar-agar are thoroughly washed and purified. The solid layers are covered with the solutions under investigation, and the relative speed of diffusion is observed. Colloids diffuse only slightly into the solid mass, but most electrolytes and many other molecular dispersoids diffuse into the gelatinous masses for a considerable distance.

The facts established by the foregoing data on diffusion, that colloids (the solid gelatin and agar-agar being nothing else but gelatinized emulsoids) are permeable to molecular dispersoids but not to other colloids, lead to the recognition of colloids as dialyzing membranes. Dialysis, therefore, is closely connected with diffusion. These data are of great importance to biologists, since they lead to the conclusion (now well established) that the colloidal dispersoids of the body cannot diffuse through the colloidal membranes of the body. For example, colloidal foodstuffs cannot get through the intestinal walls (colloidal membranes); they must first be broken down to the crystalloidal state. For further data on diffusion and dialysis, see page 235.

17. Osmotic Pressure of Colloidal Solutions.—We have seen that one of the outstanding properties of colloids is their slow diffusibility,

especially into gelatin or agar jellies. The connection between diffusion and osmotic pressure has also been emphasized. Furthermore, it has been repeatedly proved experimentally that the osmotic pressure of a solution is proportional to the number of particles of the disperse phase. In fact, all the data that we have at our command would lead us to conclude that the osmotic pressure exerted by colloidal solutions would be small, and such is the case. In some instances it is so small as to escape detection, yet we have reason to believe that all colloidal solutions exert some osmotic pressure. The experimental determination of the osmotic pressure of colloidal solutions is complicated by the difficulty of removing the last trace of electrolytes, which, owing to their great osmotic activity, obscure the actual osmotic effect of the colloid.

Many studies have been undertaken to establish laws regarding the osmotic pressure of colloidal solutions. The most elaborate work in this field is perhaps that of Lillie,[5] who studied the osmotic pressure of a number of emulsoids. He found that a solution of egg albumin containing 12.5 grams to the liter showed, at room temperature, an osmotic pressure of about 20 mm. of mercury. When the osmometer reached this height it stood constant for a considerable time. This fact was interpreted as indicating that the pressure was due to the colloid, for if it were due to some crystalloid mixed with the colloid, then on standing in contact with the membrane this would have diffused out of the colloid through the membrane, and the pressure registered on the osmometer would not have remained constant over any length of time. Several others have obtained results that agree with those of Lillie. In fact, many colloids are found to be quite stable and to manifest the same osmotic pressure for months. It is interesting to note that under the ultramicroscope these solutions show no change in the number of particles.

When we turn to suspensoids the results are less satisfactory. The final osmotic pressure seems to be largely dependent on the method of preparation of the sol. In all cases the values of the osmotic pressure of suspensoids are very small and by no means concordant.

The influence of added substances on the osmotic pressure of colloidal solutions is of considerable interest to the biologist, for, whereas in the case of molecular dispersoids this is purely additive, in the case of colloids the effect is variable, the added substance sometimes causing

[5] Am. J. Physiol., 20, 127, 1907.

an increase, and at other times causing a decrease in the osmotic pressure of the colloid. When the hydrogen-ion concentration of a solution of gelatin is raised, the osmotic pressure at first decreases, then rapidly increases. The addition of alkali tends to increase the osmotic pressure until a maximum is reached, then it begins to fall. Both acids and alkalies lessen the osmotic pressure of egg albumin. In general, electrolytes tend to decrease the osmotic pressure of such emulsoids as gelatin and albumin solutions, the extent of this influence being dependent on the nature of the cation and anion composing the electrolyte. The heavy metals exert a greater influence than the alkaline earths, and these in turn exert a greater influence than the alkalies; the *anions* arrange themselves in the order of their depressing influence as follows:

$$SO_4 > Cl > NO_2 > Br > I > CNS.$$
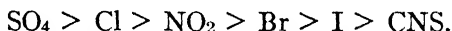
The reaction of the colloid, however, has a very great influence on the nature of the effect produced by a given electrolyte.

18. Viscosity of Colloidal Solutions.—In any liquid there is friction when the molecules rub against one another. This friction is more in evidence when a liquid flows through a tube, because the layer in immediate contact with the wall is held almost at rest, while there is a layer of varying thickness in which the velocity increases from nearly zero to a maximum. The particles of this layer, therefore, move over one another and produce friction. This internal friction, or *viscosity*, of a liquid is proportional to the velocity with which these particles are moving past one another and also to the extent of the rubbing surfaces. The *coefficient of viscosity* is the force required to move at unit velocity a plate of unit surface separated from another plate of the same size by a layer of liquid of unit thickness.[6]

When a liquid contains substances dissolved in it, or particles in suspension, the viscosity is increased. It has been found by experiment however, that the viscosity of suspensoids does not perceptibly differ from that of pure water; the viscosity of emulsoids, on the other hand, even at small concentrations, is much greater than that of water at high temperatures, and it increases rapidly on cooling. A 1 per cent solution of agar forms a jelly at room temperature. The viscosity of emulsoids also increases with increasing concentration, as is shown by the data of Table XXI.

[6] Bancroft, Applied Colloid Chemistry, 190, 1921.

TABLE XXI

VISCOSITY OF EMULSOIDS

Sol	Temperature, 20° Concentration,	Viscosity
Gelatin................	1	0.021
Gelatin................	2	0.037
Silicic acid.............	0.81	0.012
Silicic acid.............	0.99	0.016
Silicic acid.............	1.96	0 032
Silicic acid.............	3.67	0.165

Viscosity of water at 20° = 0.012.

In the living organism, the problem of viscosity reaches its greatest importance in the case of the blood. "The internal friction of the blood gives rise to what is often called the 'peripheral resistance' of the vascular system. This it is, that with a given rate and strength of heart beat, determines the arterial pressure" (Bayliss). The viscosity of serum is about twice that of pure water, while that of whole blood is six or seven times that of water. This shows that the viscosity of blood is largely due to the corpuscles. It is evident, therefore, that the viscosity is lower in defibrinated, or "laked," blood than in normal blood. Dilution also has the effect of diminishing the viscosity, so that a dilute blood will pass more rapidly through kidneys, and thus favors the excretion of urine. The relationship of the number of corpuscles to the viscosity of blood may be seen from the following figures (Dupré, Demming and Watson):

Number of Blood Corpuscles		Viscosity of Serum + Corpuscles
0	....	1.9
3,200,000	....	3.3
6,300,000	....	4.9
12,600,000	....	15.6

The figures are for horse's blood at 32.2°.

Carbon dioxide causes the corpuscles to imbibe water, thereby increasing their size and consequently the viscosity. This is the reason that the viscosity of venous blood is greater than that of arterial blood.

Viscosity of blood saturated with O_2.....................	5.4
Viscosity after H_2 is passed through it...................	5.9
Viscosity after CO_2 is passed through it.................	6.57

The intake of food causes an increase in the viscosity of blood; a meat diet increases it most, while a carbohydrate and fat diet produces a medium effect. Hunger diminishes the viscosity.

That the proteins of the plasma are also of great importance in maintaining the right viscosity of the blood is shown by the following statement of Mathews:[7]

"The great importance of the viscosity of the blood, and of the proteins of the plasma in determining it, is shown by the circumstance that if the blood is withdrawn from a dog's body and the corpuscles removed by centrifugalization and then the corpuscles suspended in a salt solution of the proper strength, such as Ringer's solution, and reinjected, the animal will not live. But it will live if gum arabic is added to the Ringer's solution in sufficient quantity to restore the viscosity to its normal amount."

19. Brownian Movement.—The invention of the ultramicroscope and the study of the phenomena exhibited by colloidal solutions when observed by this instrument have directed attention afresh to an observation which was made just about a century ago by the botanist, Robert Brown, and which has since been the subject of repeated investigations. Many microscopic objects suspended in water had been observed to be in rapid motion, and up to 1827 this phenomenon was supposed to be connected entirely with living matter. This view was disproved by Brown's experiments. He began by mixing in water pollen from ripe anthers, and he observed that the pollen appeared as very minute spherical bodies which were in rapid motion in the liquid. (The phenomenon of pollen dust had already been observed by previous investigators.) In order to ascertain whether or not this motion was a phase of life, Brown examined the spore dust of mosses that had been dry for a century. To his surprise, these particles were just as lively as the pollen he had previously examined; he therefore tested finely divided inanimate substances in water, such as gums, resins, wax, coal, glass, rocks, manganese, lead, bismuth, nickel, antimony, arsenic, and sulphur. He found the same unique motion in every case. His conclusions may best be summed up in his own words:

"Extremely minute particles of solid matter, whether obtained from organic or inorganic substances, when suspended in pure water or some other liquid, exhibited motions for which I am unable to account and which from their irregularity and seeming independence

[7] Physiological Chemistry, 4th ed., 532, 1925.

resemble in a remarkable degree the less rapid motions of the simplest animalcules of infusions. The smallest moving particles observed and which I have termed active molecules appear to be spherical or nearly so, and to be between 1/20,000 and 1/30,000 of an inch in diameter; other particles of considerably greater and various size, and either of similar or of very different figure, also present analogous motions in like circumstances.

"I have formerly stated my belief that the motions of the particles neither arose from currents in the fluid containing them, nor depended on that intestine motion which may be supposed to accompany its evaporation.

"These causes of motion, however, either singly or combined with others—as the attractions and repulsions among the particles themselves, their unstable equilibrium in the fluid in which they are suspended, their hygrometrical or capillary action, and in some cases the disengagement of voltaic matter, or minute air bubbles—have been considered by several writers as sufficiently accounting for the appearances. Some of the alleged causes here stated, with others which I have considered it unnecessary to mention, are not likely to be overlooked or to deceive observers of any experience in microscopical researches: and the insufficiency of the most important of those enumerated may, I think, be satisfactorily shown by means of a very simple experiment (referring to evaporation)."

Some forty years after the phenomenon was studied by Brown, the subject was again investigated by Wiener, Exner, and Schultze. They also tried to single out the cause of the motion from among the various explanations offered. A résumé of their work is to be found in the *Jahresberichte* of 1867, as follows:

"Chr. Wiener instituted microscopic observations of these movements, and came to the conclusion that this trembling, irregular, unsteady motion of solid molecules, which alter their direction in the briefest fraction of time in their zig-zag course, has for its basis the continual movements which, by virtue of their molecular constitution, belong to fluids. He learned through his investigations that (1) the movements are not those of infusoria; (2) the movement is not communicated to the fluid; (5) the trembling movement is not in any way derived from the varying attraction or the collisions of the various oscillating molecules; (4) the movement is not derived from changes in temperature; (5) the movement is not derived from evaporation.

Consequently, there remained, in his opinion, nothing to account for the movements but the property of the fluid itself. This explanation received direct confirmation from Wiener's observations, that the speed of the movement has a certain relation to the size of the molecule. Lately, S. Exner has extended the investigations of Wiener. Exner sought to test with reference to the molecular motion whether or not either chemical causes or mechanical ones, such as pressure, vibration, and so forth, could in any way produce an acceleration or a retardation of the effect. Only by exposure to light and heat did the motion become accelerated, and then in such a manner as, in the case of glycerine, to give freer motion to the molecules on account of decreasing the viscosity of the liquid. Exner also examined into the properties of the fluids in which solid molecules remain suspended. The results of his investigation resolve themselves into the following points: (1) the liveliness of the molecular movement is heightened by light and heat, and by radiant as well as by conducted heat; (2) one of the consequences of the molecular movement is, that the molecules, in a specially lighter fluid, not only do not sink to the bottom, but overcome the force of gravitation to such a degree as to spread themselves equally throughout the fluid; (5) the velocity of this scattering is influenced by light and heat. It should be mentioned here that Fr. Schultze had already stated that substances, when most finely divided, especially such as seemed under the microscope to be amorphous, and exhibited the Brownian movements, remain suspended in pure water and in many other fluids for days, weeks, and months at a time, so that the fluid containing them presents a cloudy or at least an opalescent appearance."

Twenty-five years later, a similar conclusion was reached independently by Gouy,[8] who showed that neither light nor convection currents within the liquid could possibly give rise to the motion. He also showed the movement to be independent of external vibration, and only slightly influenced by the nature of the suspended particles. The smaller the particles and the less viscous the suspending medium, the more rapid the motion was found to be.

Up to the advent of the ultramicroscope, the numerous investigators of Brownian movement had established the following facts: The rate of the movement is independent of the chemical nature of the particles, but depends on three factors, viz.: (1) the size of the

8 J. de Phys., 7, 561, 1888.

particle, (2) the temperature, and (3) the viscosity of the dispersion medium. The rate is increased by decrease in the mass of the particle, by increase in temperature, or by decrease in the viscosity of the medium. The movement persists, never changing, *once equilibrium has been established*, for all time. It has been observed in granite and in other rocks in small pockets of liquid, which they must have occluded for millions of years. But equilibrium must first be established. Finally the movement was explained on the grounds of collisions between the particles and the molecules of the liquid.

After Zsigmondy invented the ultramicroscope, it was found that the phenomenon becomes much more conspicuous with particles of ultramicroscopic dimensions, i.e., with colloidal particles. Zsigmondy describes his observation of colloidal gold as follows:

"A swarm of dancing gnats in a sunbeam will give one an idea of the motion of the gold particles in the hydrosol of gold. They hop, dance, spring, dash together and fly apart so rapidly that the eye can scarcely make out their movements. . . . These motions show that there is a continual mixing together of the interior parts of every liquid going on unceasingly, year in and year out."

The Brownian movement has been carefully studied in recent years by many physicists and chemists, but more especially by the French physicist, Perrin. These investigators have come to the conclusion that what we see is only an example of what is going on all the way down the line from the largest particles to the molecules themselves. They conclude that all the particles of every size, whether visible or invisible, have by their mutual impacts imparted their kinetic energy to each other until, on the average, they all have the same amount. The difference in the rate of movement is due to the fact that, to have the same kinetic energy, the large particles need only move more slowly, while the smaller particles must move more rapidly, and the smallest particles most rapidly. This supposition has been proved experimentally through the application of mathematical formulæ worked out by Einstein and Smoluchowski.

The theory of Einstein[9] in regard to the motion of particles in solution, assumes no difference between true molecules and particles suspended in the same medium. The particles behave as if they were true gas molecules with normal kinetic energy but having a much

[9] Drudes, Annalen d. Phys., **4**, 17, 549–560, 1905; **19**, 371–381, 1906; Zeit. Elektrochem., **14**, 235–239, 1908.

shorter free path. The following is the Einstein formula:

$$A = \sqrt{t}\sqrt{\frac{RT}{N} \cdot \frac{1}{3\pi\eta r}},$$

where

A is the amplitude (in the direction of the x axis);

t is the corresponding time;

R is the ordinary gas constant;

T is the absolute temperature;

N is the number of molecules in one gram-molecule;

η is the viscosity of the medium;

r is the radius of the particles (average).

Smoluchowski;[10] on totally different grounds, arrived at the same formula, except that another integer is introduced.

By the use of the above formula and also of other formulæ, Perrin calculated the value of N (Avogadro's number). The details of his experiments can be found in his book "Les Atomes." One of his methods depends on the fact that if the Brownian movement of particles is really the same as the movement of the molecules in a gas, their vertical distribution in equilibrium must follow the same law as that of the atmosphere, under the influence of gravity. In order to verify this experimentally, it was necessary to prepare suspensions of particles of uniform size and sufficiently large for observations to be made in the depth of the cell on the stage of the microscope. The object of this arrangement was to enable him to count the particles. Gamboge and mastic were the colloidal substances used. By a process of centrifugation, preparations containing particles of uniform size were made. From these experiments, a value of 70.5×10^{22} was found for the number of molecules in 22.4 liters of a gas.

Perrin also calculated the value of N by another formula of Einstein, which gives the mean displacement of a given particle in a given time in terms involving N (Avogadro's number), together with other values capable of experimental determination. The horizontal projections of the paths of three different grains in suspension of mastic are shown in Fig. 28, the dots representing the successive positions occupied by the particles at intervals of 30 seconds. By this method, a value of N was obtained that approximated the figure given above.

[10] Drudes, Annalen d. Phys., **4,** 21, 756–780, 1906; **25,** 205–226, 1908.

Perrin calculated the value of N by still another method and got an analogous figure. He gives, as the mean value of all his experiments, $N = 68.5 \times 10^{22}$.

Figure 28 will serve to give some idea of the complexity of the movements of particles in Brownian movement, but a limited one, since it must be remembered that, if the position of a particle had been mapped at more frequent intervals, a path fully as elaborate as those of the whole figure would have to be inserted between each of the positions indicated in the tracing.
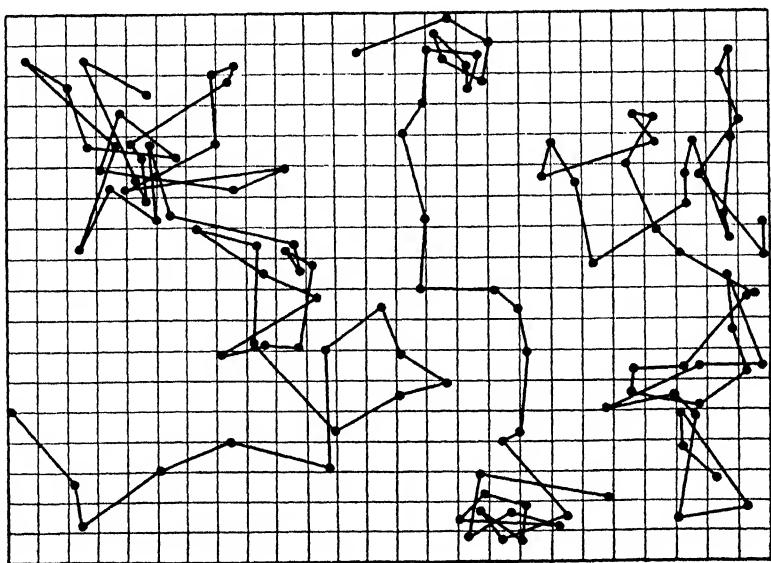
FIG. 28.

The most accurate study thus far made of Brownian movement in gases was that carried out by Millikan[11] and by Fletcher,[12] both of the University of Chicago. They employed a tiny drop of oil as the colloidal particle. Such a drop, when watched through a telescope, is seen to dart rapidly about to and fro in all directions. From a large number (5900) of measurements of the displacements along one axis of twelve such particles, Fletcher obtained the value, $N = 60.30 \times 10^{22}$ $\pm$ 1.2 per cent for Avogadro's constant. Millikan's figure of 60.62

[11] Phys. Rev., **1**, 220, 1913.
[12] Phys. Rev., **4**, 453, 1914.

$\times 10^{22} \pm 0.01$ per cent for N, which is taken as the most exact figure yet obtained, was gotten by an entirely different method.

The above experiments, together with numerous others, confirm quantitatively the theory that the Brownian movement is the result of molecular impacts. In other words, the molecules of a liquid are assumed to be in a state of ceaseless motion, and the motions of these molecules are imparted to the particles in suspension.

ELECTRICAL PROPERTIES OF COLLOIDS

20. Colloidal Particles are Electrically Charged.—When we were studying true solutions we learned that all chemical compounds could be divided into two great classes, electrolytes and non-electrolytes. The former class, in the presence of water and other dissociating solvents, break down into charged parts, or ions, the cations carrying positive charges, and the anions negative charges. The non-electrolytes are not dissociated but remain in solution in the molecular state of dispersion.

Whether the colloidal particles are charged or uncharged is just as fundamental a question for colloids as for true solutions of electrolytes and non-electrolytes.

When a difference of potential is established between two electrodes immersed in a colloidal solution, the suspended particles move with or against the current. Now, it is evident that only electrically charged particles are thus transported; hence, the colloidal particles must be charged. In fact, the charges of the colloidal particles not only determine many of their physical properties but also, as we shall see, the stability of the colloid itself.

21. Electrical Transference of Colloids (Cataphoresis).—The movement of the suspended particles through the liquids in which they are suspended, under the action of an external electromotive force, is known as *cataphoresis*. The migration of the suspended particles to the poles, is, to some extent, analogous to the migration of the ions in a true solution when under the influence of an electromotive force. This phenomenon was first critically studied by Linder and Picton.[13] They showed that when two wires connected with the terminals of a battery were placed in a colloidal solution of arsenious sulphide the compound was attracted to the positive pole, and was gradually trans-

[13] J. Chem. Soc., **61**, 148, 1892.

ferred thither. Ferric hydroxide sol was, on the other hand, attracted to the negative pole. Linder and Picton extended their study of this phenomenon and obtained the following results: Suspensions of aniline blue, arsenious sulphide, indigo, iodine, shellac, silicic acid, starch and sulphur in water, and bromine in alcohol were attracted to the positive pole. Suspensions of ferric hydroxide, hæmoglobin, Hoffmann's violet, magdala red, methyl violet, and rosaline hydrochloride were attracted towards the negative pole.

As a conclusion to their work, Linder and Picton make the following significant statement: "Experiment seems to show that, if the solution is basic or tends to break up so as to leave a free base active, the motion is to the negative pole, i.e., the particle is positively charged; if the solution is acidic, motion is to the positive pole, and consequently the particles are negatively charged."

The various colloids may be divided into those that move towards the anode, and are, hence, electronegative; and those that move to the cathode, and are, accordingly, electropositive.

ELECTRONEGATIVE	ELECTROPOSITIVE
The sulphides of arsenic, antimony, and cadmium.	The hydrates of iron, chromium, aluminium, copper, zirconium, cerium, and thorium.
Solutions of platinum, silver, gold, and mercury.	Bredig's solutions of bismuth, lead, iron, copper.
Stannic acid and silicic acid.	
Aniline blue, indigo, eosin, fuchsin, soluble Prussian blue, and molybdena blue.	Hoffmann's violet, magdala red, methyl violet, rosaniline hydrochloride.
Iodine, sulphur, selenium, shellac, resin.	Basic proteins, histones, and protamines. Proteins in acid solution.
Most natural proteins in neutral or slightly alkaline solution.	Aluminium hydrate. Hæmoglobin.
Lecithin and phosphatides.	
Gum arabic, glycogen, starch, nucleic acid, and soaps.	

The determination of the *velocities* with which the particles move in cataphoresis is not a difficult matter. The method is analogous to that employed in determining the absolute velocities with which ions move. Cotton and Mouton[14] made direct measurements of the velocity with which the particles move, by observing with a microscope the distance over which a single particle traveled in a given length of time under definite potential gradient. The simplest method of measuring the

[14] J. Chim. Phys., **4**, 365, 1906.

velocity of the movement of the particles, however, is to place the colloidal solution in the lower part of a U-tube of the type indicated in Fig. 29, which is filled up on both sides with distilled water in which the electrodes are placed. The latter may then be connected with the terminals of a battery or with the electric lighting system, and the speed of the moving boundary observed directly. If the arms of the U-tube are graduated in millimeters throughout their length, the reading of the moving boundary is facilitated. The results of a large number of experiments show that most particles move at the rate of 10–40×10^{-5} cm. per second for a potential gradient of 1 volt per centimeter. It is interesting to note that this is also the order of migration velocity of the ions. This seems to indicate that particles of all sizes—microns, submicrons, amicrons, and ions—move with approximately the same speed in the electric field.

FIG. 29.—Apparatus for Measuring Velocity Migration of Colloidal Particles. The upper part of each arm is filled with water; the colloidal solution is run in from below. (After Bayliss.)

22. Electrical Endosmosis.—The process described above may be reversed; that is, the water may be moved under the influence of electrical difference in potential, provided the suspension is held fixed. The "fixing" may be effected by placing a diaphragm or a number of capillary tubes at the bottom of the U-tube, and then filling both arms of the tube with water. If an electrode is introduced into each branch, the water will move under the influence of an electric current, and, in fact, will rise towards one of the electrodes. The diaphragm may be made of any solid or semi-solid material that is permeable to water, such as a porous clay wall, or solidified gelatin or agar.

The phenomenon of endosmosis was first observed by Reuss in 1807, and has since been studied by others, notably by Perrin,[15] who worked out a satisfactory theory of the process.

[15] J. Chim. Phys., **2**, 601, 1904.

The height to which the liquid will rise, or the difference in level of the liquid in the arms of the U-tube, is proportional to the electromotive force applied to the electrodes.

Most substances when brought in contact with water become negatively charged—the water, positively. Under these conditions water would, of course, go to the cathode. On the other hand, certain substances acquire a positive charge when brought in contact with water; in these cases the water acquires a negative charge, and hence the direction of migration is reversed.

The effect of electrolytes on electrical endosmosis has also been studied by Perrin and others. They found that acids render negative diaphragms less negative and positive diaphragms more positive. Alkalies, on the contrary, render negative diaphragms more negative and positive diaphragms less positive. Very small amounts of electrolytes exert a marked effect on electrical endosmosis. There is an important connection between the valency of the ions and the action of the salts yielding these ions on the difference of potential which exists between the liquid that is present and the solid. When the diaphragm is positively charged, the difference of potential is found to be determined by the valency of the anion. When the diaphragm is negative, the valence of the cation conditions the difference of potential.

These relations are of importance in connection with precipitation of colloids by electrolytes.

23. Action of the Electric Current on Emulsoids.—The particles of emulsoids, when acted upon by an electromotive force, are also observed to move, but the motions are much more sluggish and the phenomenon is less distinct and more complicated than is the case with suspensoids. The emulsoids exhibit a greater tendency to precipitate, and they are more sensitive to their environment, i.e., the same emulsoid particles under one set of conditions will move in one direction, and under other conditions in the opposite direction, showing that the particles may be charged either positively or negatively. This is especially true of the proteins in connection with changes in the pH of the solution.

24. Theory of the Electrical Charge of Colloidal Particles.—Many theories have been advanced to explain the origin of the electrical charges on the colloidal particles, but Noyes[16] thinks that only two

[16] J. Am. Chem. Soc., **27**, 85, 1905.

deserve consideration. He says: "In regard to the cause and character of the electrification two assumptions deserve consideration; one, that it is simply an example of contact electricity, the colloid particle assuming a charge of one sign and the surrounding water the other. This correlates the phenomena of the migration with that of electrical endosmose. It does not, however, give an obvious explanation of the facts that the basic colloidal particles become positively charged and the acidic and neutral ones negatively charged. The other assumption accounts for these facts. According to it the phenomenon is simply one of ionization. Thus each aggregate of ferric hydroxide molecules may dissociate into one or more ordinary hydroxyl ions and a residual positively charged colloidal particle, and each aggregate of silicic acid or stannic acid molecules into one or more hydrogen ions and a residual negatively charged colloidal particle. . . . To explain the behavior of neutral substances like gold or quartz by this hypothesis, it is necessary to supplement it by the assumption that in these cases it is the water or other electrolyte combined with or absorbed by the colloidal particles which undergoes ionization. It seems not improbable that there may be truth in each of these hypotheses, contact electrification occurring in the case of coarse suspensions and ionization in the case of those which approximate more nearly to colloidal solutions."

PRECIPITATION OF COLLOIDS BY ELECTROLYTES

25. General Remarks on the Action of Electrolytes.—One of the most important chapters of colloidal chemistry concerns itself with the phenomena that take place when electrolytes are added to solutions of emulsoids and suspensoids. In general, it may be said that precipitation tends to take place and that the reaction is an irreversible process. The action of electrolytes on emulsoids, however, is much more obscure then the action of electrolytes on suspensoids; hence, we shall consider first their action on the latter, and then their action on emulsoids.

All colloidal solutions are more or less unstable; the degree of instability, however, depends somewhat on the methods employed in their preparation. Electrolytically prepared metal sols are the most unstable. They degenerate very quickly, the particles clumping and settling to the bottom. Heat hastens this process, and if certain electrolytes are added the precipitation will take place at once. Some exceptions to this general relation have been noted; for instance, it

has been found that when gold from a gold sol was deposited by cataphoresis it gradually passed back again into the colloidal state, and sometimes formed a homogeneous solution. A somewhat similar observation was made by Linder and Picton[17] in connection with a ferric hydroxide hydrosol, which had been precipitated with sodium chloride. The precipitate could be restored to the colloidal state by removing the electrolyte with water. Precipitated arsenic trisulphide can be restored to the colloidal state simply by treating with hydrogen sulphide.

26. The Action of Electrolytes on Suspensoids.--When successive small quantities of a salt are added to a suspensoid sol, the solution usually becomes opalescent, and gradually a flocculent precipitate appears, which settles out, leaving a clear supernatant dispersion medium. The amounts of the various ions necessary to cause these effects, however, are very different, and many experimenters have attempted to solve the underlying principles.

The fact that gold hydrosols are precipitated by salts was demonstrated by Faraday in 1858,[18] and it was this action of salts that stimulated other investigators to pursue the subject. Schultze[19] established the fact that the power of various electrolytes was greatly increased by valency, but to an extent much beyond the relation obtainable by the number of electrical charges. Hardy[20] used more quantitative methods, and established a law according to which, if the precipitating power of a monovalent ion is represented by x, that of a divalent ion will be x^2, and that of a trivalent one, x^3. This law is only an approximation, but it helps to give one an idea of the difference in the precipitating power of the ions of higher valency. For instance, it has been found that the precipitating power of an aluminium ion, Al^{+++}, is hundreds of times as great as that of a sodium ion, Na^+, towards a negative colloid, whereas we might have expected it to be only three times as great. Hardy also showed, in the same article, that the precipitating ion is always one whose charge is of the opposite sign to that of the colloid precipitated. He formulated the following rule: "The coagulative power of a salt is determined by the valency of one of its ions. This prepotent ion is either the negative or the positive

[17] J. Chem. Soc., **61**, 114, 1892; **87**, 1924, 1905.
[18] Phil. Trans., **147**, 145, 1858.
[19] J. pract. Chem., **25**, 431, 1882.
[20] J. Phys. Chem., **4**, 235, 1900.

ion according to whether the colloidal particles move down or up the potential gradient. The coagulating ion is always of the opposite electrical sign to the particle." This is known as *"Hardy's rule."* The numbers given in Table XXII give the concentration (in millimoles, or thousandths of a gram-molecule per liter) of different electrolytes, required to effect the precipitation of an arsenious sulphide sol, illustrate the rule. The figures are taken from the experiments of Freundlich,[21] who showed that an amount of electrolyte which is incapable of bringing about the precipitation of the colloid instantaneously, may become effective after an interval of time. He also showed that the total quantity of electrolyte required to precipitate a suspensoid completely depends somewhat on whether the electrolyte is added all at one time or in successive portions. He consequently adopted the following procedure, which tends to prevent the possibility of irregularities due to the time factor: One cubic centimeter of electrolyte solution was added to 10 c.c. of the sol, the mixture shaken and allowed to stand for 2 hours.

TABLE XXII

Precipitating Action of Electrolytes on Arsenious Sulphide Sol

(7.54 millimoles As_2S_3 per liter.)

Electrolyte	Concentration (millimoles per liter)
Monovalent cations:	
KCl	49.5
KNO_3	50.0
LiCl	58.0
NaCl	51.0
HCl	30.8
Divalent cations:	
$BaCl_2$	0.69
$MgCl_2$	0.72
$MgSO_4$	0.81
$CaCl_2$	0.65
$Ba(NO_3)_2$	0.69
$ZnSO_4$	0.81
Trivalent cations:	
$AlCl_3$	0.093
$Al(NO_3)_3$	0.095

[21] Zeit. phys. Chem., **44**, 131, 1903.

It will be seen from the examples cited that a greater amount of salts containing monovalent metals (cations) is required to precipitate an electronegative colloid than is required of salts containing divalent metals; and a greater amount of these is required than is required of salts of trivalent metals.

The same general rule has been observed in connection with the effect of the anion of a salt in the precipitation of electropositive colloids.

The precipitating efficiency of an ion does not depend entirely on its valency but is also dependent on its chemical nature, as is evidenced by the fact that ions of the same valency vary considerably in their efficiency. The readily adsorbed H^+ and OH^- ions, which, as we have seen, have a powerful effect in electroendosmosis, have a greater precipitating power than other monovalent ions. In fact, any ion that is readily adsorbed seems to have a lower precipitation concentration than other ions of the same valency. It is evident, therefore, that the maximum precipitation of a colloid is not necessarily coincident with the point of absence of charge, i.e., the iso-electric point; yet this is what we should expect, if the electrical attraction of opposite charges were the only force to be considered. For example, the precipitation of sulphur is effected by an amount of cæsium which is about one hundredth of the amount of lithium required to produce the same result. Now, it certainly does not require one hundred times more lithium, to neutralize the charge, than cæsium. In short, there is abundant evidence to show that the phenomenon is a complicated process involving the neutralization of the electrical charge (and, hence, affected by the valency of the precipitating ion), the specific adsorption of the ion, the chemical nature of the reacting particles, and the stabilizing effect of the ions carrying the same charge as the colloid.

27. Explanation of the Valency Rule.—Various explanations have been advanced to explain the valency rule. Qualitatively, there is little difficulty in explaining this phenomenon if it is remembered that precipitation is a consequence of removing a definite charge from the surface of the colloidal particles by means of the ions of the opposite charge. That there may be *unneutralized charges within the colloidal particle* is evidenced by the fact that a smaller amount of the precipitating salt is required when it is added quickly and in bulk. It is also evident that, since the charges on monovalent, divalent, and trivalent ions are in the ratio 1 : 2 : 3, the divalent ions have a greater

chance of neutralizing the surface charges on the colloid particles than the monovalent, and the trivalent than the divalent. In the words of Hardy: "Since it is the valence of the number of electrical charges on the ion, which is of importance in precipitating, there is a far greater chance of two charges arriving simultaneously in the neighborhood of a colloidal particle when both charges are on the same ion, than when they are on separate ions. For a trivalent ion the chances are very much better that the three should arrive simultaneously if all are on the same ion than if each is on a separate ion." Mathews[22] offers the following explanation: "Bivalent and trivalent ions are more effective in precipitating because they unite two or three or more colloidal aggregates into very large aggregates of the following kind: Ca—colloid —Ca—colloid—Ca—colloid—Ca—colloid."

28. The Stabilizing Effect of Ions of the Same Charge.—The number of charges carried by the ions of the same charge as the colloid is of little importance, since they do not unite with the colloid; nevertheless, these ions exert some effect on the precipitating particles. Take the following examples: The precipitation concentrations for ferric hydroxide are 9.64 for $BaCl_2$, 400 for HCl, and 0.42 for $Ba(OH)_2$; the precipitation concentration of sodium chloride on a platinum sol was 2.5; of sodium hydroxide, 130; the figures for a mastic were $NaCl$ = 0.12 and HCl = 0.004.

It would seem, therefore, that the precipitation concentration of a given electrolyte is the result of a balance between the stabilizing effect of the ion of the same charge as the colloid, and the precipitating effect of the ion of the opposite sign.   This is one of the reasons why the valency rule is not mathematically exact.

29. Adsorption of the Precipitating Ion.—One of the most instructive discoveries in connection with the precipitation of colloids by salts is that when a colloid is precipitated the precipitating ion is always found in the precipitate. Thus, when antimony sulphide is precipitated by sodium chloride, there is always some sodium in the precipitate; when it is precipitated by barium chloride, there is barium in the precipitate, and so on. Furthermore, the supernatant liquid is found to be acid. It has been found next to impossible to remove these ions by washing with pure water, but they can be removed by washing with a solution of some other salt, the cation of the solution taking the place of the original precipitating ion. The substitution is quan-

[22] Physiological Chemistry, p. 235.   Wm. Wood & Co., New York, 1925.

titative. If the precipitating ion is an anion the same results are obtained.

Lange[23] has shown that the precipitation of suspensoids has an important diagnostic application in connection with the examination of cerebro-spinal fluid of patients suffering from diseases of the central nervous system (e.g., general paralysis of the insane, syphilis, meningitis): "When gold sol is added to 0.4 per cent solutions of NaCl containing varying concentrations of cerebro-spinal fluid, the concentration at which precipitation of the colloidal gold occurs differs according to the general character of the disease."

30. Precipitation of Emulsoids by Electrolytes.—Much less is known regarding the action of salts on emulsoids than on suspensoids, yet the emulsoids are much more important from a biological standpoint than the suspensoids. It will be recalled that the emulsoids include colloidal solutions of proteins, starch, glycogen, dextrins, agar, soaps, and silicic acid. The action of electrolytes on emulsoids is much more obscure than the action of electrolytes on suspensoids. Usually, large amounts of electrolytes are necessary to effect precipitation, and Hardy's rule does not seem to manifest itself. This may be due in part to the fact that emulsoids occupy a position intermediate between suspensoids and true solution, and in consequence posses a greater degree of dispersity. Some emulsoids approach the suspensoids in their properties, and others true solutions. It is evident, then, that the differences between suspensoids and emulsoids are merely a matter of degree. If we keep in mind the fact that emulsoid sols are hydrophilic (Greek, "water-loving") particles suspended in a solution of the same substance, it will not be difficult to comprehend the difference between the properties of these sols and those of suspensoid sols. Physical chemists express the distinctive property of emulsoid sols by stating that they are *highly solvated*. Many emulsoid sols are very viscous, and become more viscous with decreasing temperature, finally settling to a stiff jelly. A suspensoid sol will not form a jelly, no matter how concentrated it is made

The precipitation of a suspensoid is usually an irreversible process, i.e., the precipitate will not pass back into the colloidal state. But when an emulsoid, such as albumin, is thrown down (this usually requires a very large amount of a salt; e.g., if ammonium sulphate is used, the solution must be nearly saturated before precipitation takes

place) it will usually redissolve if placed back in distilled water and washed free from adsorbed electrolytes. Some suspensoid precipitates, however, are reversible; e.g., a precipitated gold suspensoid can be redissolved by dilute ammonia; while precipitated emulsoids, on the other hand (albumin), enter into the irreversible state on standing, which prevents them from being redissolved when treated with water. It should also be noted that while relatively concentrated solutions of the alkalies and alkaline earths are required to precipitate such emulsoids as albumin, very dilute solutions of the salts of the heavy metals are found to be sufficient to bring about complete precipitation.

PROTECTION AND MUTUAL PRECIPITATION OF COLLOIDS

31. Protective Action of Colloids.—The typical emulsoids, as we have seen, are usually very stable and not very easily precipitated by electrolytes. The suspensoids, on the other hand, are readily precipitated by electrolytes. If one adds an emulsoid to a suspensoid colloid, the particles of the latter become coated, by adsorption, with the emulsoid, and the stable characteristics of the emulsoid are transferred to the suspensoid. In this manner, precipitation of suspensoids by electrolytes is greatly lessened, if not almost entirely prevented. The emulsoids are, accordingly, said to exert a *protective action* on the suspensoids, which are thereby rendered more stable. This is another of the numerous facts that were discovered by Faraday while studying the properties of colloidal gold. He discovered that the addition of a trace of "jelly" to his gold solutions made them much less sensitive to precipitation by salt.

There is considerable difference in the protective power of different electrolytes. This fact was shown very conclusively by Zsigmondy[24] while studying the action of various emulsoids on colloidal gold. A solution of red colloidal gold becomes blue on the addition of a small amount of sodium chloride, owing to the increase in the size of the particles. Various emulsoids, when added to a red gold sol, protect the colloidal particles from precipitation by a solution of sodium chloride, so that no change of color follows the addition of the electrolyte. Zsigmondy's method of studying the effects of the protective action of the various substances was as follows: He took a certain amount of a solution of colloidal gold and added varying amounts of

[24] Zeit. analyt. Chem., **40**, 697, 1901.

emulsions and emulsoids and 1 c.c. of 10 per cent sodium chloride solution. The amount of colloid, in milligrams, which just failed to prevent the change from red to blue, he termed the "*gold number*." A few of Zsigmondy's results are given in Table XXIII.

TABLE XXIII

Colloid	Gold Number
Gelatin....	0.05 -0.01
Casein..........................	0 01
Egg albumin..........	0.06 0 3
Gum arabic....	0 15 0.25
Dextrin.	6- 12
Starch (wheat)..	4 6
Starch (potato)....	25

Colloidal solutions of the sulphides are protected, as well as the colloidal metals; and silicic acid, as well as organic colloids, exerts a protecting influence on the metal sols.

This stabilizing action of emulsoids is of considerable importance in the industries and in mary physiological processes. A small amount of gelatine is often added to commercial ice cream to keep the mixture smooth by preventing the growth of crystals. The gelatine in the ice cream also aids digestion by preventing the separation of the fats and the coagulation of the milk in heavy curds. Gum arabic is added to marshmallows, also for the purpose of preventing crystallization. Colloidal lubricants and fuels are stabilized with tannin or rosin soaps. Glue is added to the chemical mixtures used to treat boiler water to prevent the formation of scale. It is supposed to act by coating the precipitated particles, thereby preventing them from adhering to the boiler tubes. It has been found that when sparingly soluble salts of calcium, e.g., calcium phosphate, are formed in the blood or serum, very much larger amounts than those corresponding to the solubility of the salts can remain in colloidal suspension. The deposition of gall stones (consisting of cholesterol and the calcium salt of bilirubin, for example) has been traced to a reduction in the amount of protective colloids in the bile, brought about by pathological conditions. In the normal bile, the bile salts act as protective colloids which keep the various sparingly soluble substances in a state of colloidal suspension. If these protective colloids are reduced in quantity, however, separation of the sparingly soluble substances takes place.

In general, then, it may be said that when an emulsoid is mixed with a suspensoid sol in certain proportions, the suspensoid sol acquires most of the characteristics of the protecting colloids: a positive suspensoid will lose its positive character, if protected by an electronegative emulsoid. The masking of the properties of a suspensoid sol by a protecting colloid is, therefore, to be ascribed to the formation of a thin film of adsorbed emulsoid over the suspensoid.

32. Mutual Precipitation of Colloids.—Since an electrically charged colloid can be precipitated by an ion of the opposite sign, it is natural to presume that an oppositely charged colloid would have the same effect. Many investigators have worked with this problem, but Blitz[25] was the first to give a satisfactory explanation. He demonstrated that when a solution of a positively charged colloid is added to a negatively charged colloid, precipitation occurs, but when colloids of the same electrical charge are mixed no precipitation takes place. In the case of oppositely charged colloids, there is a definite concentration at which both colloids are precipitated quantitatively. This is indicative of a definite chemical reaction.

When we were considering the precipitation of colloids by electrolytes, we saw that the amount of precipitation was conditioned not only by the amount of electrolyte added to the colloid, but also by the way in which it was added, whether slowly or all at once. The same relation obtains for the precipitation of a colloid with one of the opposite sign. If, to a given sol, another sol of opposite charge is added more quickly one time than another, there will be a difference in the amount of precipitation. Furthermore, if a large amount of the colloid of opposite sign is added at one time, precipitation may not take place. Likewise, if only a small amount of the colloid of opposite charge (an insufficient amount to neutralize the charge) no precipitation is likely to take place. Evidently, the colloid, which is present in excess, acts as a protective colloid in these cases. Some of the results obtained by Blitz (*loc. cit.*) are given in Table XXIV which shows that to obtain a maximum precipitating effect a definite relative concentration must be maintained, and that if one deviates too far, in one direction or the other, from this relation, there will be no precipitation. To 2 c.c. of antimony trisulphide sol, containing 2.8 mgs. per cubic centimeter, the varying amounts of ferric hydroxide sol were added.

[25] Ber. **37**, 1095, 1904.

TABLE XXIV

Milligrams Fe₂O₃	Results directly after Mixing	Results after One Hour
20.8	cloudy.............	homogeneous
12.8	cloudy.............	slight ppt.
8.0	slight ppt...........	complete ppt.
6.4	complete ppt........	complete ppt.
4.8	flakes..............	liquid yellow
3.2	small flakes.........	unchanged
0.8	cloudy.............	nearly homogeneous

When emulsoids are added to suspensoid sols the general reaction is that described under the caption of "protective colloids." Some precipitation make take place if they are oppositely charged, but the reaction is never complete, and as a rule the suspensoid acquires the charge of the emulsoid owing to the coating effect. No generalizations have been established.

When we undertake the study of the action of one emulsoid on another, we find that the same general conditions that were observed when oppositely charged electrolytes were added to emulsoids obtain also in this instance. Some precipitation does take place with the oppositely charged emulsoids, but the close resemblance between emulsoids and true solutions renders the phenomenon more or less indefinite. When the two colloids are mixed in equivalent electrical quantities, more or less complete precipitation occurs; in any other relationship, the colloid that is present in excess exerts a protective action on the other. The phenomenon, as a whole, is very complex, and no general relations have thus far been established.

THE PHENOMENA OF SURFACE TENSION AND ADSORPTION

33. Surface Tension.—The molecules in the interior of a homogeneous liquid do not exhibit any surface phenomena in relation to one another, since they are subject to a *balanced* attraction, i.e., the molecules of a liquid of this type are attracted to one another in all directions, and these attractions neutralize one another. But the surface film of a liquid is in a state of tension due to the *unbalanced* attractions between the molecules at the surface. The surface molecules are attracted only downward and sideways, not upward, since

there are no liquid molecules above them. It follows that at the surface of liquids there is a force—the so-called *surface tension*—acting inwards, the molecules behaving as if they formed a tightly stretched but elastic skin over the surface of the liquid. Fine powders, of such a nature that they do not readily take up moisture (such as sulphur) float when sprinkled on water; the particles rest on the surface exactly as if supported by a membrane. Substances that tend to lower the surface tension—bile salts, for instance—are dissolved in the water, and hence will not be buoyed up, but will sink. Surface tension is manifested also at all points where the liquid comes in contact with the container, or any solid substance placed within the liquid.

The conditions obtaining in colloidal solutions, however, are quite different; the same surface relations are exhibited at the surface and at the points of contact with the supporting vessel, but, in addition, the suspended colloidal particles are large enough to separate the molecules of the solvent beyond the range of their cohesional attractions. In other words, the colloidal particles possess surface, and hence the phenomena of surface tension are exhibited at the interfaces between the colloidal suspension and the solvent.

Some liquids wet the walls of a glass capillary tube, while others do not. When a liquid wets the walls of the capillary tube, the liquid not only rises in the tube, but assumes a concave shape; on the other hand, when a liquid does not wet the surface of the capillary tube, the surface is depressed and assumes a convex shape. Likewise, if a glass plate is placed in a liquid that wets it, the surface of the glass will be moist for some distance above the surface of the liquid. This amounts to an increase in the total surface. Now, the force of surface tension operates against an increase in surface, and tends to pull the surface into the least possible area. This it does by raising the liquid at the points of junction of the surface of the glass; hence, the surface curves upward on the glass. The area of the surface along this curved portion is considerably less than the area of the surface in this vicinity if it were lying flat plus the area of the vertical film of liquid on the glass plate. The curved portion of the surface is called the meniscus.

34. Determination of the Surface Tension.—The surface tension of a liquid can be determined in several ways. Only a few of these will be considered here. One of the most accurate is the so-called ripple method of Lord Rayleigh, which consists in measuring the speed of propagation of a series of ripples set up in a pan of the liquid. A

mathematical relation exists between the velocity and the surface tension. This method is reliable with pure liquids only.

Another method is based on the principle discussed above, i.e., on the rise or fall of the level of a liquid in a capillary tube, according to whether it wets the glass or not. This method offers an easy means of estimating the surface tension of solutions by comparing the surface tension of the solution with that of the pure solvent. For example, the relative surface tension of a solution may be determined by measuring the height of the column of liquid in a capillary tube, and comparing with the height of the column of the solvent when the same tube is used. The specific gravity of the solution must also be measured, so that the comparison made will be of weights of liquids raised. If the exact diameter of the bore of the capillary is known, the absolute surface tension can be calculated.

The surface tension may also be determined by the "drop method." If one slowly removes the finger from a pipette containing a liquid, one will observe that when a drop forms at the tip, it appears to be exactly spherical while small, but on growing it assumes a bag-like shape. As soon as the drop has increased in weight a sufficient amount to overcome the tension which has been holding it up, it will fall. The larger the drop, the greater the surface tension. The number of drops that are formed from a given volume of liquid in a given time is counted by means of an instrument designed for the purpose, called a stalagmometer (Fig. 30). The

FIG. 30.—A = Stalagmometer; B = Bottle.

specific gravity of the liquid must be taken into account, and a correction must be made for temperature. Since it has been found that the surface tension diminishes very rapidly with rise in temperature, all determinations must be subjected to a correction, or, better still, the determinations should be carried out at a constant temperature. In order that this may be easily done, Findlay has suggested passing the end of the stalagmometer through a rubber stopper, which in turn is placed in the neck of a small bottle

(Fig. 30). This apparatus may then be placed in a thermostat or constant-temperature bath.[26]

The formula for the calculation of the surface tension, in dynes, is as follows:

$$\text{Dynes per cm.} = \frac{\text{Number of drops of water}}{\text{Number of drops of liquid}} \times \text{sp. gr.} \times 73.$$

The factor 73 is the surface tension of pure water in dynes, determined for a line at the surface of the water 1 cm. in length.

The values of the surface tensions of pure liquids show great variation. The following numbers, in dynes per centimeter, will serve to illustrate this: At 20° C. the surface tension of water is 73; of ethyl alcohol, 22; and of ethyl ether, 16.

35. Surface Tension of Solutions.—Most salts increase the surface tension of the solution, the increase being very nearly proportional to the concentration. The strong bases increase the surface tension, but ammonia lowers it, as do also the strong mineral acids. The surface tension between liquid and air is lowered when organic substances are dissolved in the liquid. The degree of lowering varies considerably according to the organic substance added. Some substances, bile salts, for example, have a very great effect. The same statement applies to the interface between liquid and liquid, except that in this case all substances in solution, even inorganic salts, *lower* the surface tension.

36. Surface Tension at Various Interfaces.—The phenomenon of surface tension is not confined to the surface of liquids; i.e., it is not confined to the interface between various liquids and air. As a matter of fact, the phenomenon is exhibited wherever the liquid comes in contact with the containing vessel, with suspended solid particles, or with other liquids with which it does not mix.

In the case of the immiscible liquids, the interfacial surface will be curved and will have its concave towards the side of greater surface tension, and the liquid on this side will tend to form drops in the other liquid. If hydrated particles (emulsoid colloids) are present in suspension in water, there is also a liquid-liquid interface about each particle.

Interfaces between liquids and between suspended solid particles and liquids are met with in biology much more frequently than those between gases and liquids.

[26] Findlay, Practical Physical Chemistry.

37. Surface Tension of Colloidal Solutions.—From careful measurements it has been found that coarse suspensions and suspensoids produce very little effect on the air-liquid surface tension of the dispersion medium (water); emulsions and emulsoids, on the other hand, decrease the air-liquid surface tension of the dispersion medium to a considerable degree. Such solutions froth profusely when they are shaken. Frothing is due to a lowering of the surface tension of the dispersion medium by the suspended particles. In this case, lowering of the surface tension takes place at the interfaces of the two liquid phases; i.e., the interfacial tension is lowered, hence, the tearing of the films of the continuous phase between the particles is prevented. Since suspensoid solutions do not froth to any extent, this difference in behavior can be used as a means of distinguishing between emulsoids and suspensoids.

Very little is known about the conditions of stability for emulsoids; it can be shown, however, that by lowering the surface tension of two *immiscible* liquids against each other, an emulsion is obtained. When one shakes water and oil together vigorously, the oil is broken up into very fine drops, but no permanent emulsion is formed. If, however, a few drops of potassium or sodium hydroxide are added, a milky emulsion is formed on shaking, and does not show any noticeable change, even after many hours. The emulsifying action is attributable to the lowering of the surface tension of water by the soap, and to the formation of a film on the surface of the droplets, the coalescence of which is thereby prevented. It can, furthermore, be shown that permanent emulsion and emulsoid droplets tend to acquire an electrical charge and hence mutually repel one another.

The decrease in the liquid-liquid surface tension of emulsions can be measured by means of a special pipette invented for the purpose by Donnan. This pipette allows a definite volume of liquid A to run into liquid B. The pipette is provided with a stopcock and a capillary outflow, bent upward. The lower the surface tension between the two liquids, the greater the number of drops. Therefore, by counting the number of drops in a given time, one obtains an approximate measure of the surface tension of the emulsion or emulsoid.

38. Gibbs-Thomson Principle.—From the above data it is evident that the addition of a solute to a solvent induces a change in the attractive force between the molecules of the solvent, for this force varies with each molecular species. In other words, the addition of a

solute necessitates a redistribution of energy. Willard Gibbs[27] has formulated a theorem which gives a mathematical expression to the phenomenon. Quite independently, though at a later date, Sir J. J. Thomson[28] promulgated the same principle. The Gibbs-Thomson principle states that *substances which lower the surface tension become concentrated in the surface layer, whereas the concentration of those substances which raise the surface tension becomes less in the surface layer than in the body of the solution.* This principle is of fundamental importance in the field of biochemistry, since the phenomenon known as *adsorption* depends on it. Practically all the distinctive properties of colloidal solutions, e.g., electrical changes, precipitation, Brownian movement, viscosity, and imbibition, are more or less connected with adsorption phenomena.

39. General Remarks on Adsorption.—It has long been known that charcoal has the power to take up certain gases, like carbon dioxide, in large quantities. It can also remove certain constituents from solution; for instance, if water containing a coloring matter, such as caramel or litmus, is shaken up with finely divided charcoal, the latter on settling carries down the coloring matter with it, leaving the water practically colorless. Further investigation has shown that charcoal will take up other substances from solution besides gases and coloring matter, e.g., electrolytes and non-electrolytes, and even colloids.

Charcoal is not alone in possessing this adsorbing power. Practically all substances in a finely divided state, and, therefore, exposing large surfaces, have the same property. Solids not in a finely divided state will also adsorb substances from solution, but only to a limited extent, owing to the relatively limited surface exposed. *Adsorption*, then, is the tendency exhibited by all solids to condense upon their surfaces a layer of any gas or liquid with which they may be in contact. The term was invented to distinguish the phenomenon from "absorption," which means the accumulation of one substance within the interior of another.

Surface adsorption is a common phenomenon. Gibbs (*loc. cit.*) showed that adsorption will take place whenever the surface tension of the adsorbing body is lowered by the concentration in the surface layer of the material in the solution or other surrounding medium

[27] Scientific Papers, I, 219.
[28] Thomson, Applications of Thermodynamics to Physics and Chemistry.

(see above). A great many cases are known where there is apparently neither chemical reaction between the adsorbing and the dissolved substances nor penetration of the former by the latter. Many cases are known in which both adsorption and absorption take place; adsorption, however, usually takes place much more rapidly than the penetration into the interior.

As applied to colloidal phenomena, adsorption may be exhibited in many different ways. A few examples will be given: (1) Dissolved crystalloidal substances, both electrolytes and non-electrolytes, may be adsorbed on the colloidal particles of a hydrosol, and in consequence will not dialyze or react chemically to the extent that they would in ordinary solution. (2) In the process of gel formation, electrolytes and other soluble substances are largely adsorbed out of the solution and appear in the gel. For this reason it is difficult to obtain absolutely pure gels. (3) Colloidal substances are adsorbed by mechanical suspensions like charcoal, fuller's earth, etc. This fact has been utilized to the greatest advantage in such processes as the purification of sugar. In the case of sugar, the charcoal removes from solution the colloidal proteins, coloring matter, etc., with which it is contaminated. (4) Colloids mutually adsorb each other, as in the case of the "protective colloids" which were previously considered. (5) The first stage in the process of dyeing and staining is one of adsorption. At a later stage solid solution and chemical reaction may play a part. All of these types of adsorption are of interest to the physiologist, on account of the part they play in biological phenomena.

40. Adsorption of Gases.—All solids tend to adsorb any gases with which they are in contact. The amount of adsorption for given conditions of pressure and temperature will, of course, vary with the nature and physical condition of the solid and with the nature of the gas. But in all cases the process of adsorption of a gas is a reversible one; i.e., as soon as the process of adsorption has been initiated it will continue until a condition of *adsorption equilibrium* has been attained.

When one determines the amount of gas taken up by a given sample of charcoal[29] at a constant temperature, it is found that the adsorp-

---

[29] One charcoal is not like another. The adsorption results vary greatly according to the wood from which it is made, and charcoal from the same wood will vary with the method of preparation. Cocoanut charcoal has the greatest adsorbing power of all, and hence is the most widely used in adsorption experiments.

tion, when the pressure or concentration of the gas is varied, is relatively much greater at lower than at higher pressures. These facts can be observed by reference to the results obtained by Travers[30] with carbon dioxide and charcoal. The adsorption of carbon dioxide by charcoal at 0° C. is given in Table XXV.

TABLE XXV

Pressure in mm. (P)	x Concentration of CO_2 in charcoal	$\dfrac{x}{\sqrt[3]{P}}$
5.1	0 38	0 24
25.1	0 77	0.26
137.4	1.45	0.26
416.4	2.02	0.27
858.6	2 48	0.26

From the numbers in the third column, it is seen that the concentration of carbon dioxide in the charcoal is proportional to the cube root of the pressure, or the cube root of the concentration of the gas.

The amount of gas adsorbed is found to be proportional to the adsorbing surface, and is greater the lower the temperature as well as the higher the pressure. Titoff[31] tested the adsorption of hydrogen, and found that it follows Henry's law; therefore the formula

$$C_A/C_B = \text{constant}$$

applies, where C_A represents the concentration in the solid phase, and C_B that in the gas phase. Other gases do not follow Henry's law so closely. The adsorptive power of charcoal for traces of gas, especially at low temperatures, has been used by Dewar to produce a high vacuum; the pressures were too low to be capable of measurement.

Advantage is taken of the adsorptive power of finely divided solids for gases in the use of charcoal in "poison gas" masks and for the removal of foul gases in many other instances. Since the amount of adsorption is reduced by rise in temperature, it is evident that these processes are most effective at low temperatures.

[30] Proc. Roy. Soc., A, 78, p. 9, 1906.
[31] Zeit. phys. Chem., 74, 641, 1910.

Contact catalysis can also be explained by the great adsorptive power of charcoal (see p. 326).

41. Adsorption from Solution.—As has already been stated, Gibbs, from theoretical considerations, arrived at the conclusion that when the surface tension of a liquid is lowered by a solute, the solute will become concentrated in the surface layer of the solution. This theory has been proved again and again. It is evident, then, that if we increase the extent of the boundary layer by introducing into the solution a solid dispersoid possessing a relatively large surface area, e.g., a very fine powder, a highly porous substance, or a colloid, the dissolved substance will tend to become concentrated in the layer of solution in contact with the solid. After the adsorption of the solute on the surface of the solid, a new equilibrium will be established. If the original concentration of solute is not too great and sufficient solid is added, the concentration of solute in the solution may approach zero as a limit, especially if it has a great effect on the lowering of the surface tension of the liquid.

As was the case with gases, relatively more material is adsorbed out of dilute solutions than out of concentrated ones. A ten-fold increase in the concentration of the dissolved material results in only four times as much adsorption by the suspension or colloidal substance which may be introduced into the two solutions. In this, adsorption differs from chemical action, which is proportional to the concentration of the reacting substances present in a solution.

The adsorption out of different liquids by the same adsorbing body is different in amount. For example, from aqueous solutions charcoal adsorbs acids, bases, and salts only slightly; while organic substances, especially the aromatic compounds, are strongly adsorbed. From solutions in organic solvents, the dissolved substances are only slightly adsorbed. The following data will illustrate this relation; the concentration of benzoic acid in the several solvents is kept constant and is equal to 0.01 mole per liter.

	Amount adsorbed in millimoles
Benzoic acid in water........................	3.27
Benzoic acid in benzene......................	0.54
Benzoic acid in acetone......................	0.30
Benzoic acid in ether........................	0.29

A general relation has been worked out between the adsorption of a substance dissolved in any given solvent and the adsorption of that

solvent when dissolved in any other solvent: "Solvents from which the dissolved substances are easily adsorbed, are, when dissolved in other solvents, only slightly adsorbed. The converse also holds: substances which from solutions are strongly adsorbed, when used as solvents yield solutions from which the dissolved substances are only slightly adsorbed. Thus, sulphuric acid is, as we have seen, only slightly adsorbed from aqueous solutions. Therefore, substances dissolved in this solvent should show strong adsorption, and such is the case."[32]

As a general rule, adsorption takes place very rapidly. In other words, the equilibrium between the adsorbed substance and the adsorbent is quickly reached. The following figures for succinic acid (the adsorbed substance) and charcoal (the adsorbent) will confirm this conclusion. The amounts of acid are expressed in milli-equivalents per gram of charcoal.

Minutes	Amount adsorbed
5	0.183
10	0.325
30	0.752
60	1.060
∞	1.160

It is evident that adsorption is practically complete at the end of an hour.

The effect of raising the temperature on adsorption is interesting, since it is just the opposite of the effect on solution. Most substances dissolve much more rapidly at higher than at lower temperatures, and in far greater amounts. In a few instances, however, exactly the reverse is true.

The effect of raising the temperature on the adsorption of acetic acid from an aqueous solution by charcoal may be observed from the following data:

Temperatures	Adsorption
0°	1.150
50.2°	0.728
93.8°	0.466

Another noteworthy fact, concerning the adsorption equilibria between carbon and an aqueous solution of some substance that can

[32] Jones, Nature of Solutions, p. 295, 1917.

be adsorbed, is that the process is reversible. If a definite weight of
an adsorbent is agitated with a given quantity of solution of known
concentration, a definite equilibrium will, of course, be established.
On diluting this mixture with a known volume of solvent, the adsorp-
tion will decrease until it acquires the same value it would have
attained, provided the same weight of adsorbent had been originally
placed in the more dilute solution. An illustration of this important
point may be quoted.[33] After 1 gram of charcoal had been agitated
with 100 c.c. of a 0.0688 M solution of acetic acid for 20 hours, adsorp-
tion was found to have reduced the original concentration of the acid
to 0.0608 M. In a second experiment, 1 gram of the same charcoal
was shaken, for a similar period of time with 50 c.c. of a 0.1376 M
($= 0.0688 \times 2$) solution of acetic acid, and then, after 50 c.c. of water had
been added, the shaking was continued for an additional 3 hours. The
final concentration of the acid was found to be almost identical with
that of the first experiment.

42. The Adsorption Isotherm.—When a solid adsorbent, such as
charcoal, is agitated with a solution containing an adsorbable sub-
stance, the amount of substance adsorbed by a given mass of the
adsorbent depends entirely on the concentration of the solution and on
the temperature. Another way of stating this fact, is as follows: For
a given concentration and temperature, there is a definite amount of
adsorption, which varies only with the nature of the adsorbent and the
solution. This *adsorption equilibrium* has been expressed by Freund-
lich (*loc. cit.*) in the form of an equation known as the *adsorption
isotherm*. The equation may be written as follows:

$$\frac{x}{m} = k \cdot c^{\frac{1}{n}}.$$

where x is the weight of substance adsorbed by a given weight, m, of
adsorbent, from a solution whose volume concentration at equilibrium
is c, and k and n are constants. k is the amount of adsorption that
takes place when the concentration is equal to unity, and it is peculiar
to the adsorbed substance. The constant n varies only within narrow
limits.

Table XXVI contains the data given by Freundlich[34] on the

[33] Freundlich, Zeit. phys. Chem., **57**, 385, 1906.
[34] Kapillarchemie, Leipzig, 1909.

adsorption of acetic acid from solutions of different strength by charcoal.

TABLE XXVI

ADSORPTION OF AQUEOUS ACETIC ACID BY CHARCOAL

$t = 25°; \ k = 2.606; \ 1/n = 0.425$

Concentration in moles per liter	x/m (obs.)	x/m (calc.)
0.0181	0.467	0.474
0.0309	0.624	0.596
0.0616	0.801	0.798
0.1259	1.110	1.080
0.2677	1.550	1.490
0.4711	2.040	1.890
0.8817	2.480	2.470
2.7850	3.760	4.010

A satisfactory agreement between the experimentally determined and the calculated values of x/m is shown by the above figures.
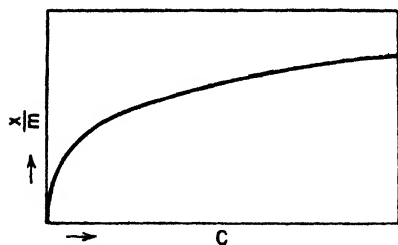
FIG. 31.—Adsorption Isotherm

It is customary to express the data given in Table XXVI in the form of a curve, known as the isothermal adsorption curve, or *adsorption isotherm* (see Fig. 31). It will be noted that, when we plot x/m against c, that is, adsorption against the equilibrium concentration, the curve becomes a parabola (or approximately so).

It is noteworthy that the adsorption-equilibria phenomena, which are obtained with acetic acid and charcoal in aqueous solution, are found to obtain also with a number of dyes. If we take a series of solutions of a dye, such as Congo red, in progressively diminishing concentration, and place in each solution the same amount of filter paper, it will be found that a part of the dye is taken up with the paper, and the relative amount taken up is larger the more dilute the solution. This relation was extensively studied by Bayliss, and is referred to by him as the law of adsorption. Table XXVII will illustrate this point:

TABLE XXVII

Concentration of Solution		Proportion of Dye in Solution	Proportion of Dye in Paper
Initial	Final	Per cent	Per cent
0.014	0.0056	40	60
0.012	0.0024	20	80
0.010	0.0009	9.3	90.7
0.008	0.0003	4	96
0.006	0.00008	1 3	98.7
0 004	......	trace	practically all
0.002	......	trace	practically all

This behavior of dyes is of importance, not only in the process of dyeing commercial fabrics, but also in connection with the staining of tissues in histology.

43. Theories of Adsorption.—When gases, liquids, and dissolved particles are adsorbed by solids, they tend to disappear as such and are condensed on the surface of the solids. Surface adsorption is a very common phenomenon. As we have seen, it was extensively studied by the physicist, Willard Gibbs, who showed that adsorption will take place whenever the surface tension of the adsorbing body will be lowered by the concentration in the surface layer of the material which is available in the solution or other surrounding medium.

Since substances in the colloidal state have very large relative surface areas, it follows that the opportunity for surface adsorption on colloidal substances is very great. (The extent of surface in a colloidal solution is enormous in proportion to the mass of substance in solution. For example, a 10-c.c. sphere with a surface of 22 sq. cm., if reduced to a fine powder consisting of spherules 0.00000025 cm. in diameter, will have a surface of 20,000,000 sq. cm., i.e., nearly half an acre.) The amount of material adsorbed by a given colloid depends, however, not only on the extent of surface, but also on (1) the chemical nature of the particle, (2) the electrical charge on the surface, (3) the temperature, and (4) the concentration of the material adsorbed.

The nature of adsorption is still a matter of discussion. This is due to the fact that the rules which apply in some cases do not hold in others. Thus, non-electrolytes and electrolytes that reduce the surface tension are adsorbed, but neutral salts, such as sodium chloride, that

do not lower the tension of the water-air surface are nevertheless adsorbed by charcoal, paper, etc.

The electrical charge of the substance adsorbed and the charge on the adsorbent do, however, play important parts in most adsorption processes. We have seen (p. 260) that, when an ion is adsorbed by a colloid in the process of precipitation, the ion cannot be removed by washing with neutral water, but that it can be replaced by washing with a solution containing an ion carrying the same charge. From this fact and other data, Langmuir has suggested that, in some cases at least, adsorption is simply a chemical union of a layer of solute molecules with the surface layer of molecules of the substance on which it is adsorbed (adsorbent). In many cases, however, this seems not to be true. McClendon has suggested that, in the case of many organic compounds, large molecules that are not particularly attracted by the adsorbent, upon being driven out of solution, will be adsorbed because they have no other place to go.

Another important aspect which must not be forgotten is that the solvent is adsorbed on a surface in addition to, or even in preference to, the solute. From some salt solutions, for example, water may be adsorbed in greater proportion than the salt. This is sometimes called "negative adsorption" from the point of view of the solute, since its concentration is lower on the surface than in the body of the solution and the latter becomes more concentrated.

44. Reactions that Depend on Adsorption.—Adsorption plays a large part in biological phenomena as well as in industrial procedures. In fact, it would be hard to find a chemical process in which adsorption does not play some part.

Many industrial processes are based upon adsorption phenomena. Of course, large numbers of these processes were known and practiced long before the phenomenon itself was even partially understood. But with the coming of the knowledge of the nature, causes, and possibilities of the control of the phenomenon, immense improvements have been effected. Any discussion of the industrial applications of adsorption phenomena would be out of place here; but the following list of examples will serve to illustrate the immense importance our knowledge of adsorption plays in industry: the contact sulphuric acid process; the Hargreaves salt-cake process; the Deacon chlorine process; the Chance-Claus process of making sodium carbonate; the Haber ammonia process; the Ostwald nitric acid process; the Sabatier

hydrogenation process; the manufacture of soap; the dyeing of fabrics; the tanning of leather, etc. This list could be extended to great length, but it is already long enough to emphasize the importance and practical value of the phenomenon in daily life.

The following are a few instances in which adsorption undoubtedly plays a very important part in biological processes:

(a) *The Action of Enzymes.*—This topic will come up for discussion later (see p. 333). At present it is only necessary to state that all biological catalysts, or enzymes, are colloidal in nature, and hence, as would be expected, adsorption enters into the process of their activity.

(b) *The Action of Toxins and Anti-toxins.*—The neutralization of toxins by anti-toxins also occurs according to the laws of adsorption.

(c) *Adsorption Compounds.*—There are certain complex compounds which appear on investigation to be adsorption compounds rather than chemical compounds. *Jecorin* is such a complex. This substance was originally obtained from the liver, but has since been made artificially from lecithin, a protein, and a carbohydrate, by mixing these substances in varying proportions. Mixtures of proteins and inorganic salts are also adsorption compounds. In fact, there is considerable evidence to show that adsorption compounds are very abundant in living cells, and that they are constantly being formed and broken down.

(d) *The Action of Drugs.*—Adsorption plays a very important part in the action of certain drugs which do not enter the cell, but are adsorbed on its surface and act from that situation. Thus, the action of corrosive sublimate (mercuric chloride) on microörganisms has been found by Morawitz to be proportional to the amount adsorbed on their surfaces. It has also been found that strophanthin itself does not enter the cells of the frog's heart.

(e) *Purification by Adsorption.*—The purification of water by filtration through beds of sand depends largely on the removal of impurities of all types by adsorption on the large filtering surface exposed. It is interesting to note, in this connection, that Indians sometimes filter strongly alkaline water through fine silt to remove the excess of alkali. Some cacti contain a large amount of reserve water in a spongy mass in their centers. This water is quite salty and bitter, but when no other water is obtainable the Indians filter it through fine silt, which adsorbs the excess of salts and makes it palatable.

The above are only a few of the important phenomena in which adsorption plays an active part.

<center>EMULSIONS</center>

45. Nature of Emulsions.—Emulsions are systems of two liquids insoluble in each other; they consist of comparatively coarse liquid particles of one liquid suspended in another liquid with which it does not mix. The conditions necessary to produce a stable emulsion have been concisely expressed by Hildebrand:[35] "When two incompletely miscible liquids are mechanically agitated so as to disperse one of them in the other in the form of droplets, an amount of work must be performed upon the system which is equal to the product of the interfacial tension by the increase in surface. This work may be considerable, and accounts for the fact that emulsions made from two pure liquids are always unstable, the coalescence of the droplets upon contact liberating the stored energy. To stabilize an emulsion, therefore, it is necessary to add some third substance which is capable of producing a film which will prevent the coalescence of the drops. The conditions for the stability of such a film are, in part, the same as those which give stability to the liquid film between the bubbles of a foam, or *emulsified gas.*"

Emulsions are of two kinds: (1) a small quantity of a liquid in suspension in a large amount of another liquid; and (2) a large amount of one liquid suspended in another liquid. In the second class the continuous phase must consist of a solution of a colloid, such as soap, protein, or saponin.

46. Formation of Emulsions.—

(*a*) An emulsion of oil in water is obtained if a stream of oil is injected into water. An emulsion is also formed if an alcoholic solution of oil is poured into water.

(*b*) Permanent emulsions are obtained when *colloids* are present in a solution and the solution is shaken up with another liquid. The most typical permanent emulsions are observed when fats and oils are shaken up with water containing some colloid, or the ingredients to make one.

The fats are neutral substances, but generally they contain a little fatty acid, which gives them an acid reaction and causes the formation

[35] Bogue, Colloidal Behavior, Vol. I, p. 212, McGraw-Hill Book Co., 1924.

of an emulsion when they are shaken up with an alkali. The following experiment illustrates this point:

In five test-tubes the following mixtures are placed:

(1) 10 c.c. of water + 2 c.c. of neutral cottonseed oil.
(2) 10 c.c. of water + 1 drop of 2N NaOH sol. + 2 c.c. of neutral cottonseed oil.
(3) 10 c.c. of water + 2 drops of oleic acid + 2 c.c. of neutral cottonseed oil.
(4) 10 c.c. of water + 1 drop of 2 N NaOH + 2 drops of oleic acid + 2 c.c. of neutral cotonseed oil.
(5) 10 c.c. of water + 1 drop of 2 N NaOH + 2 c.c. of commercial olive oil.

Each tube is shaken thoroughly for at least five minutes. It will be observed that only in (4) and (5) is a permanent emulsion formed, separation occurring in (1), (2), and (3) after a time. (5) shows that ordinary fat or fatty oils may contain free fatty acids.

The formation of an emulsion is due to the fact that a layer of soap, formed by the combination of the free fatty acid with alkali, is made around the fat particle.

47. Properties of Emulsions.—The properties of the first kind of emulsions (small quantities of particles suspended in a large amount of dispersion medium, the amount of suspended particles being usually about 1 in 10,000) are almost the same as those of suspensoids. The globules show Brownian movement; they are precipitated or coagulated by electrolytes and can be retained by ultrafilters. The particles are comparatively rigid and are separated from one another by thick films or layers of the continuous phase.

The properties of the second type of emulsions, in which the quantity of the disperse phase is large, are very different. They are very *viscous*, an extreme example being an emulsion consisting of 99 per cent of oil and 1 per cent of soap solution, which is of such consistency that it can be cut into cubes, i.e., it is almost a solid. Another case is cited by Pickering,[36] in which viscous emulsions were made with large amounts of kerosene in water containing small amounts of soap. With 70–80 per cent of kerosene, the mass was of the consistency of heavy cream; and with 99 per cent of kerosene, the emulsion was as stiff as blanc mange and could be cut into cubes which would hold their shape. This emulsion was not permanent, however, for when a portion was left standing in dry air it liquefied because the water evaporated from the soap film, which then cracked, allowing the kerosene to run out.

[36] J. Chem. Soc., **91**, 2002, 1902.

In these emulsions the continuous phase must be a solution of an emulsoid. Good emulsifying agents are those colloids that tend to concentrate at the liquid-liquid interface by strong adsorption. Such solutions froth when they are shaken. "Frothing is due to a lowering of the surface tension by the substance in solution. This lowering of the surface tension takes place at the points of contact between the two phases, i.e., the interfacial tension is lowered, which prevents tearing of the films of the continuous phase between the particles." (Plummer.)

48. Nature of Foams.—There is a remarkable resemblance between emulsions and foam. The former is a dispersion of liquid in liquid, and the latter a dispersion of gas in liquid. Adsorption films surround the droplets in an emulsion, and adsorption films surround the gas bubbles in foam.

The saponins make the best froth films, better even than the alkali soaps. The superiority of saponins over soaps as frothers is due to the fact that they do not lower the surface tension as much as the soaps. Owing to the tendency of soaps to lower the surface tension, they tend to concentrate at the interfaces.

EMULSOIDS

49. Distinctive Characteristics of Emulsoids.—The most important class of colloids, from the biologist's standpoint, is undoubtedly the *emulsoids*, which include suspensions, in water, of proteins, starch, dextrin, glycogen, agar, gelatin, soaps, and silicic acid. The properties of emulsoids show many differences among themselves, and differ in many ways from the properties of suspensoids and emulsions.

The emulsoids resemble the emulsions in being liquid-liquid systems and in possessing a high viscosity, but they differ in the ease with which the solvent may pass from one phase into the other. Gelatin sol is a continuous liquid phase in which droplets of high concentration are suspended; gelatin gel is a continuous solid phase containing droplets of dilute gelatin sol. When gelatin gel (jelly) is dissolved, sol formation takes place by absorption of water and swelling, thereby disintegrating the original system.

The distinction between emulsoids and suspensoids is easily observed by means of an ultramicroscope. The suspensoids are non-deformable particles, usually solid, which show active Brownian movement and are readily detectable by means of an ultramicroscope.

When emulsoid sols are examined, however, by means of the ultra-microscope, one rarely observes concrete particles, but only a general diffuse light. There are, nevertheless, many reasons for assuming the presence of submicronic particles. The failure to detect them by means of the ultramicroscope is said to be due to the fact that, being simply more concentrated droplets of the same composition as the continuous phase, they have a similarity of refractive indices, and hence there is no optical contrast between the two phases.

Another difference between emulsoids and suspensoids is their reaction towards electrolytes. We have already learned that suspensoids are precipitated on the addition of small amounts of electrolytes carrying the opposite charge (p. 257), whereas emulsoids are precipitated only by the addition of large amounts of salts and are reversible.

The word "emulsoid" indicates the liquid nature of the dispersed phase, but, since this phase may contain a greater or less percentage of water or other solvent with the same composition of the actual solid matter itself (this is especially seen in the proteins), it is evident that all degrees between solid and liquid may exist.

Emulsions are systems of two liquids insoluble in each other, whereas the properties of emulsoids indicate that they are systems of two liquid phases of different concentrations; i.e., they consist of more or less dilute solutions of a colloid containing droplets of more concentrated solution.

Another property of emulsoids of biological importance is the protective action which they exhibit against the precipitating influence of electrolytes on suspensoids (see p. 262).

50. Action of Heat on Emulsoids.—The effect of heat on emulsoids varies greatly according to the nature of the colloid, the kind and amount of electrolytes present, and especially according to the reaction of the mixture. When an albumin hydrosol is gradually heated, a temperature is finally reached at which coagulation occurs. The exact nature of this transformation is not understood, but there is evidence to indicate that it is largely chemical. This belief is based upon the fact that the formation of the coagulum is accompanied by the disappearance of the free acid from the solution, indicating the formation of a protein salt. If too much acid is present the coagulation may not take place; a moderate amount of acidity will tend to produce partial coagulation; yet slight acidity of the sol is essential to complete

coagulation. The addition of neutral salts to a slightly acid solution of a salt-free protein has been found to exert a marked influence on the temperature of coagulation. This is due in part to the decrease in the concentration of the free acid, and also in part to the direct influence which the salts exert upon the protein.

The effect of heat on a gelatin sol is very different from its effect on a protein sol. Both gelatin and agar dissolve in hot water. The solution on cooling sets to a jelly. If a rather concentrated gelatin sol is heated and then permited to cool, the jelly is not reconverted into a sol when the temperature is raised again, unless more water is added. A less concentrated jelly can be dissolved again by simple warming. The change, however, does not take place at a definite temperature. The setting temperature differs from the dissolving temperature, and both of these vary with the concentration: a 5 per cent gelatin sol melts at 26.1° and solidifies at 17.8°, while a 15 per cent sol melts at 29.4° and solidifies at 25.5°. Furthermore, the temperature of the gel-sol formation is affected by the presence of salts, some tending to raise the temperature of gelation, and others to lower it. Thus, the temperature of gelation is raised by the following anions in the order of effectiveness indicated:

sulphate > citrate > tartrate > acetate.

The following anions tend to lower the temperature of gelation:

$Cl < ClO_3 < NO_3 < Br < I < $ thiocyanate.

From these data we observe that the sulphate ion is most effective in bringing about gelation, whereas the thiocyanate ion is the most effective in retarding gelation.

Some emulsoids, for example, caseinogen and gum arabic, do not coagulate on heating, and do not form gels; the solutions simply vary in viscosity at different temperatures.

51. Structure and Physico-chemical Properties of Proteins.—The most important emulsoid colloids in the field of biochemistry are the proteins. The investigation of their behavior is rendered difficult through the complexity of their structure. It is clear, however, that the physical chemistry and biophysics of the proteins must in the last analysis depend upon their structure.

The elementary composition of the proteins was determined many

years ago, but it soon developed that this knowledge was not sufficient to establish their structure and properties. It was not until it was discovered that proteins are composed of smaller compounds, namely, amino acids, that any real light was thrown on the subject. In other words, the physical and chemical properties of proteins are now known to be dependent primarily upon the presence in the gross molecule of a preponderating amount of amino acids, each of which contains at least one basic group (NH_2) and at least one acid group (COOH). The various amino acids are linked together in a proteins by the COOH of one uniting with the NH_2 of another, with the elimination of water, thus, $R - CO \boxed{OH + H} N H - R$; but some NH_2 and COOH groups are left uncombined. It is the relative number of these uncombined radicals in polypeptids and proteins that determines whether their aqueous solutions will be faintly acid, neutral, or faintly basic in reaction; and it is the same free acid or basic groups which enable them to combine with acids or bases, or with each other, to form the vast particles upon which the colloidal properties of the proteins depend.

We have already learned (p. 193) that simple amino acids are typical amphoteric electrolytes; i.e., they react either as acids or bases. the same statement holds true for proteins, owing to the presence of the amino and carboxyl groups. This fact may be expressed by the symbol

$$\left(P \begin{array}{l} -NH_2 \\ -COOH \end{array} \right),$$
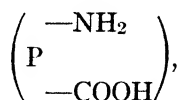
where P represents the remainder of the large protein molecule. According to the relative number of these uncombined radicals, the protein will exhibit faintly acid, neutral, or basic properties when it is dissolved in neutral water. With acids, for example, a salt will be formed by union with the NH_2 groups, which will dissociate into the anion of the acid and a large complex cation; whereas with alkalies union will take place with the COOH group, and the salt on dissociating will form a small cation of the metal of the salt and a large complex anion. In other words, the particles of the protein will have a positive or negative charge according to the chemical nature of the fluid in which it is dissolved.

If an acid like HCl is added to a protein, a hydrochloride salt is formed as follows:

$$\begin{bmatrix} & -NH_2-HCl \\ P & \\ & -COOH \end{bmatrix} \rightleftarrows \begin{bmatrix} & -NH_3{}^+ + Cl^- \\ P & \\ & -COOH) \end{bmatrix}$$
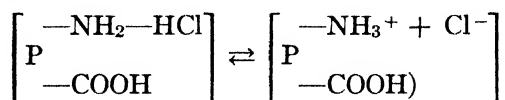
If a strong alkali, like NaOH, is added, the sodium salt is formed thus:

$$\begin{bmatrix} & -NH_2 \\ P & \\ & -COONa \end{bmatrix} \rightleftarrows \begin{bmatrix} & -NH_2 \\ P & \\ & -COO^-) + Na^+ \end{bmatrix}$$
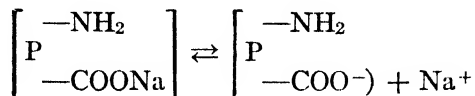
52. The Iso-electric Point of Proteins.—For every protein, as for every other amphoteric electrolyte, there is a definite concentration of H^+ ions at which the electrical conductivity is at a minimum. At this point, which is known as the "iso-electric point" for the amphoteric electrolyte (see p. 193), either the protein will not migrate in an electric field, or half will go to the anode, and half to the cathode. At this point the protein is least soluble, because it is least hydrated.

It is well to keep in mind the fact that water itself is a true amphoteric electrolyte, since it dissociates both (H^+) and (OH^-) ions. At room temperature the concentrations of (H^+) and (OH^-) in pure water are almost exactly 1×10^{-7}, i.e., pH 7 is the iso-electric point for pure water at room temperature (approximately 22° C.) It rarely happens, however, that the iso-electric point of a protein is the same as that of water; therefore, we almost invariably have to add a small amount of acid or alkali to an aqueous solution of protein to bring the mixture to its iso-electric point.

The determination of the iso-electric point was first made by electrophoresis: the solution of protein was placed in a U-tube through which an electric current was caused to flow. It was observed that the protein migrated either to the anode or to the cathode according to the reaction of the medium. At a certain pH the protein ceased to migrate. This was called the iso-electric point. For example, Michaelis and others have shown in cataphoresis experiments that gelatin particles will not migrate to one pole or the other when the pH is 4.7, whereas at any pH either above or below this point they are definitely charged and migrate to the anode in the former case and to the cathode in the latter.

Other methods which are used to determine the iso-electric point

of proteins are based on the fact that certain physical properties, such as viscosity and osmotic pressure, are at a minimum at the iso-electric point.

The iso-electric points of most animal proteins are on the acid side of aqueous neutrality (pH 7), whereas those of vegetable proteins cover a much wider range. The following are the iso-electric points of some common proteins:

Protein	pH at the Iso-electric point
Glutenin	4.45
Serum albumin	4 7
Casein	4.7
Gelatin	4.7
Egg albumin	4 8
Fibrinogen	5.0
Serum globulin	5.4
Oxyhæmoglobin	6.74
Gliadin	9.0

The iso-electric point is found to be a factor of great biological importance, especially in relation to the biophysics of the colloids found in living matter. For example, the physical properties of proteins, such as swelling, osmotic pressure, viscosity, etc., undergo decided change at this point. The curves given in Fig. 32 illustrate these various phenomena. The experiments on which they are based have been described as follows by Loeb in his book, "Proteins and the Theory of Colloidal Behavior" (New York and London, 1924), and elsewhere.[37]

"A number of doses, each containing 1 gram of finely powdered Cooper's gelatin which had a pH of a little over 7.0 and consisted partly of calcium gelatinate, were put for 30 minutes at 15° into beakers containing 100 c.c. of HBr of different concentrations, varying from M/8 to M/8192; and, as a control, 1 gram of gelatin was put for 30 minutes at 15° into 100 c.c. of distilled water. The powdered gelatin was then put into a cylindrical funnel and the acid allowed to drain off. The powdered gelatin in the funnel was then perfused six or eight times with constant stirring, each time with 25 c.c. of cold water —i.e., water not above 5°—to remove the excess of acid and the salts. The water must be cold to prevent the powdered granules from coalescing, since otherwise the washing would be incomplete. After

[37] J. Gen. Physiol., Vol. I, p. 363, 1918.

the liquid was drained off from the filter, the volume (i.e., the relative swelling of the gelatin) was measured; then the gelatin was melted by heating to 45°, and enough water was added to bring the volume in each case to 100 c.c. Then the conductivity, osmotic pressure, and viscosity were measured . . . and the pH was also determined, either colorimetrically (which gives fairly accurate results with gelatin, but not with the other proteins), or, preferably, with the hydrogen electrode. In the experiment represented in Fig. 32 the pH was measured colorimetrically. A glance at the figure shows that the ordinates of the curves representing the values for osmotic pressure, conductivity, swelling, etc., drop very sharply at pH 4.7, i.e., the iso-electric point of gelatin. By this method the approximate location of the iso-electric point can be recognized at a glance from the osmotic pressure measurements, the conductivity measurements, etc. The lowest curve in Fig. 32 represents titration for Br gelatin
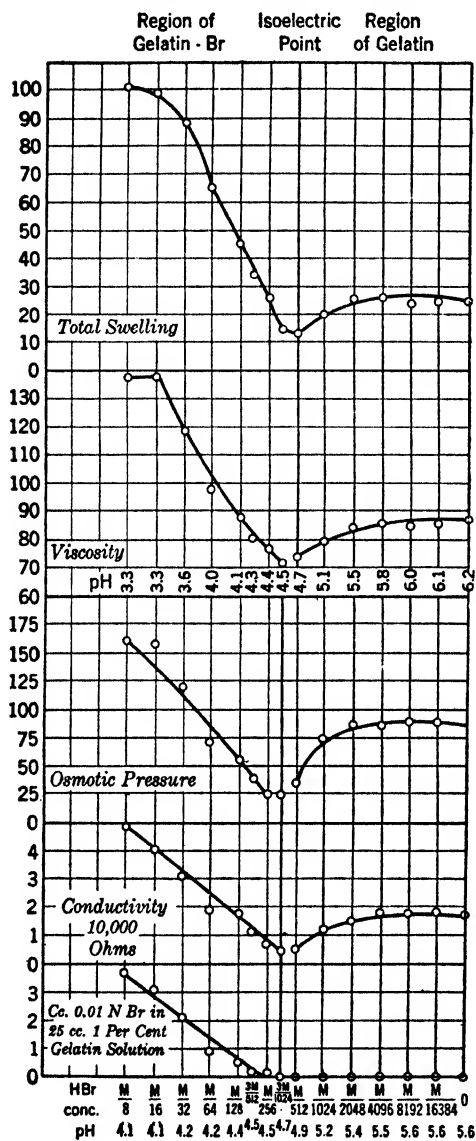
FIG. 32.—Showing that the Physical Properties of Gelatin are a Minimum at the Isoelectric Point. (After Loeb.)

should exist in the form of gelatin bromide only on the acid side

of the iso-electric point, and titration should be negative when the
pH is above 4.7. The curve shows no Br was found when the pH
was equal to or greater than 4.7; while it was found on the acid side
increasing in quantity the lower the pH. On the alkaline side of the
iso-electric point the gelatin existed still in the state of calcium gelat-
inate. In this experiment the mass of the gelatin was diminished by
solution and washing to 0.8 gram, or possibly a little less."

There are several other properties of proteins which are minimal at
their iso-electric points: (a) All proteins show their greatest tendency
to precipitate or coagulate at their iso-electric point. Even the "water-
soluble" proteins, like the albumins and gelatin, which are quite
soluble at the point of electrical neutrality, are more readily precip-
itated at this point if some precipitating agent, such as alcohol, is
added. The same proteins a little above or below the iso-electric point
may withstand the addition of a considerable excess of alcohol without
being precipitated. (b) The globulins are usually defined as being
insoluble in distilled water, but soluble in neutral salt solutions. To
this definition one should append the phrase "in the neighborhood of
their iso-electric points." When a globulin is dissolved by the addition
of a small amount of acid or alkali to the water, it will migrate in an
electric field; it will not migrate, however, if it is dissolved in a salt
solution in the neighborhood of its iso-electric point.[38] (c) It is also
interesting to note that crystallizable proteins, like hæmoglobin, will
crystallize most readily in the neighborhood of their iso-electric
points.[39]

Determinations of the iso-electric points of various proteins were
carried out by Michaelis and his co-workers several years ago, and are
summarized by Michaelis in his monograph on hydrogen-ion concen-
tration.[40] Michaelis has also enunciated an understandable exposition
of the ionization relations of amphoteric electrolytes, and a new
definition of the iso-electric point, which according to him is the point
at which the undissociated fraction is at a maximum.

**53. The Influence of Electrolytes on the Properties of the
Proteins.**—Hofmeister[41] was the first to investigate the effects of
different salts on the physical and chemical properties of the proteins.

[38] H. Chick, Biochem. J., **7**, 318, 1913.
[39] Ferry, J. Biol. Chem., **57**, 819, 1923.
[40] Michaelis, L., Die Wasserstoffionenkonzentration. Berlin, 1914.
[41] Files of the Arch. exp. Path. u. Pharm. from 1888 to 1891.

He and his followers found that the relative effects of anions on the precipitation, swelling, and other properties of proteins are very definite, and that the anions may be arranged in series according to their relative efficiency, the order being relatively independent of the nature of the cation. Such a series may be illustrated by the following sequence of ions, which are arranged according to their depressing effect on the osmotic pressure of gelatin:

$$Cl > SO_4 > NO_3 > Br > I > CNS.$$
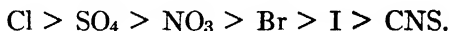
According to Loeb, however, there is a weak point in the earlier investigations in that they failed to take into account the varying effect of the salts on the hydrogen-ion concentration of the solution. He has demonstrated that the pH is a most important factor and that the salts in the complex Hofmeister series merely act as buffers, changing and stabilizing the pH. Loeb proved this by adjusting the pH of the saline solutions before the comparison was made. Under these conditions the Hofmeister series can be replaced by a simple valency rule. Acetates, nitrates, and chlorides have the same influence, from a quantitative standpoint, on the swelling, osmotic pressure, and viscosity of protein solutions, but not the same influence as sulphates and other divalent ions. Similar conditions obtain with cations. The Hofmeister series is, therefore, according to Loeb, meaningless. It is worthy of note that some authors still believe in the reliability of the Hofmeister series; however, they do not deny the validity of Loeb's work. To quote Michaelis: "Without doubt Loeb has proved the truth of his opinion in all conditions studied by him, . . . in other conditions the Hofmeister series still becomes manifest, but that does not hold in Loeb's experimental conditions."

54. Chemical Reactions of Proteins with Acids and Bases.—The chemistry of proteins does not differ from the chemistry of crystalloids, for it has been shown that the proteins combine stoichiometrically (i.e., by the purely chemical forces of primary valency) with acids and alkalies, forming protein salts which dissociate electrolytically. As in most chemical procedures, the easiest method of proving that proteins combine in simple molecular proportions is by titration experiments. The "end point," in the case of proteins, however, is different from that generally taken in analytical chemistry. In the usual chemical work, titration is carried to the point of neutrality, i.e., the iso-electric point of water—pH 7. Proteins, we have seen, are

amphoteric electrolytes having iso-electric points of their own. Gelatin and casein, for instance, act as bases at any pH below 4.7; hence, if we wish to ascertain how much of a given acid 1 gram of iso-electric gelatin or casein can bind, we have to titrate to a pH below 4.7.

Another consideration of great importance in relation to titration experiments with proteins is the discovery that the resulting hydrogen-ion concentration of the solution, when different quantities of 0.1 N acid, e.g., HCl, are added to the same quantity of protein solutions, is different from the pH which is found when the same amount of acid is added to the same quantity of pure water. This is due to the fact that part of the acid combines with the NH_2 groups of the protein molecule in the same way as if it were added to NH_3, thus forming a salt of the type $R—NH_3Cl$. Protein chloride may therefore be expected to dissociate in the following way:

$$\text{Protein—}NH_3Cl = \text{protein—}NH_3{}^+ + Cl^-.$$
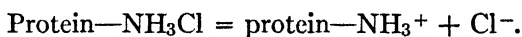
It is especially worthy of note that Loeb has demonstrated by volumetric analysis that the quantity of anion or cation in combination with a protein is, to use his own phrase, "an unequivocal function of the pH." When HCl is added to iso-electric protein and the pH is determined, the same amount of Cl is always found in combination with a given mass of originally iso-electric protein for the same pH; so that, knowing the pH and the concentration of the originally· iso-electric protein present, one can also tell how much Cl is in combination with the protein for this pH. The same is true when one adds an alkali to iso-electric proteins. For the same pH, the amount of cation in combination is always the same. These facts led Loeb to propose the following theory: "When we add an acid, e.g., HCl, to iso-electric protein and equilibrium is established between free HCl, protein chloride, and non-ionogenic or iso-electric protein; when we add alkali an equilibrium is established between metal proteinate, non-ionized protein, and the hydrogen ions."

The fact that the hydrogen-ion concentration of a protein solution determines the quantity of protein salt formed is the basis on which the experiments were devised which gave absolute proof that proteins combine stoichiometrically with acids and alkalies. We have seen that proteins combine with acids only at a pH below that of the iso-electric point, which for gelatin or casein is about 4.7, and for crystalline egg albumin 4.8. It happens that at a pH below 4.7 most of the dibasic

and tribasic acids dissociate as monobasic acids. Thus, H_3PO_4 dissociates into H^+ and the monovalent anion $H_2PO_4^-$. Hence, if acids combine stoichiometrically with iso-electric protein, it should require exactly three times as many cubic centimeters of 0.1 N H_3PO_4 to bring a 1 per cent solution of iso-electric protein to a given pH, e.g., 3.0, as are required in the case of HN_3O or HCl; while twice as many cubic
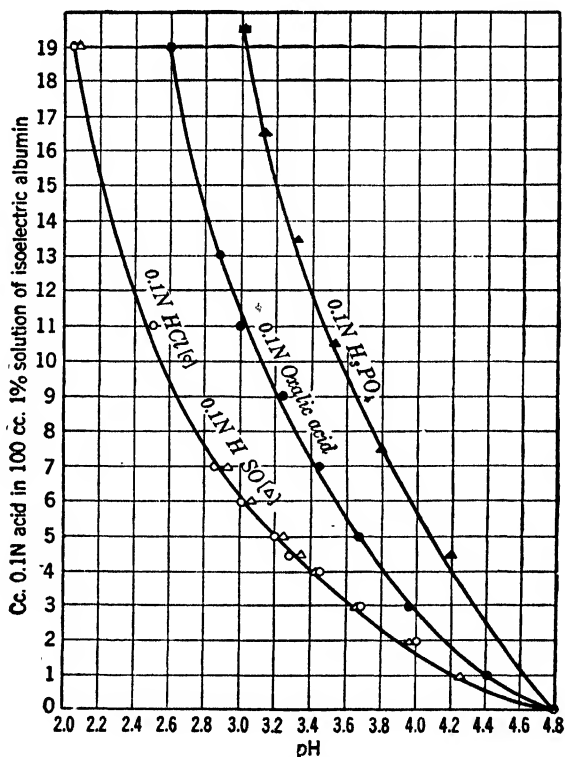
FIG. 33.—Showing Combining Quantities of Acids with Iso-electric Proteins. (After Loeb.)

centimeters of 0.1 N oxalic acid as HCl should be required. On the other hand, it should require just as many cubic centimeters of a divalent strong acid like H_2SO_4 as of HCl. A glance at Fig. 33 shows that this is the case. The ordinates of this figure are the cubic centimeters of 0.1 N acid required to bring 1 gram of iso-electric crystalline albumin to the pH indicated by the four acids mentioned, namely, HCl, H_2SO_4, oxalic, and phosphoric acids. Enough water was added to bring the albumin and acid to a volume of 100 c.c. For the same

pH the ordinates for HCl, H_2SO_4, and H_3PO_4 are approximately as 1 : 1 : 3. The ratio of HCl to oxalic is a little less than 1 : 2 when the pH is > 3.0.

On pages 294 and 295 are given the results obtained for the two proteins, crystalline egg albumin and gelatin, using HCl, H_2SO_4, $C_2O_4H_2$, and H_3PO_4, in four series of experiments for each protein.

The results show the actual numbers of cubic centimeters of 0.1 N solutions of each of the four acids in combination with 1 gram of originally iso-electric protein in 100 c.c. of solution. The values for HCl and H_2SO_4 are remarkably close. Those for H_3PO_4 are within the limits of the accuracy of the measurements, being practically three times as large as those for HCl. Oxalic acid combines with the proteins in practically twice the amount required with HCl or H_2SO_4 at the same pH. The best results, however, are below pH 3.2.

On the basis of the same reasoning that was applied to acids, one would expect that equal numbers of cubic centimeters of 0.1 N $Ca(OH)_2$ and $Ba(OH)_2$ as of LiOH, NaOH and KOH should be required to bring 100 c.c. of a 1 per cent solution of iso-electric protein to the same pH. Loeb proved that this is also the case.

These results furnish undisputable proof that the ratios in which ions combine with proteins are identical with the ratios in which the same ions combine with crystalloids, i.e., on the basic of a simple stoichiometrical combination. Interesting confirmation of this view is furnished in the recent experiments of Hitchcock[42] and Harris.[43]

55. Why are Proteins Termed Colloids?—If proteins behave like crystalloids in their chemical aspects, the question may be asked: Why are proteins termed colloids? Before answering this question it will be necessary to refer to a topic which will be presented in greater detail later, namely, Donnan's theory of membrane equilibria. This theory underlies modern progress in the physical chemistry of the proteins just as much as the idea that proteins are amphoteric electrolytes. In 1911, Donnan[44] propounded a theory of membrane-equilibria, which is now classical. He considered the equilibrium condition which results when two solutions of electrolytes are separated by a membrane impermeable to one ion of one of the solutions, but freely permeable to all other ions present. On a basis of thermo-

[42] Physiol. Reviews, **4**, 505, 1924.
[43] Proc. Roy. Soc., B. **97**, 364, 1925.
[44] Zeit. Elektrochem., **17**, 572, 1911.

dynamics he deduced certain equations connecting the ion concentrations in the two solutions at equilibrium, and predicted that there must be a difference of electrical potential between the two solutions at equilibrium, the magnitude of which must be determined by the difference in the concentrations of any diffusible ion in the two solutions. This theory was thoroughly tested by Donnan and his co-workers with a number of inorganic systems. Donnan, however, pointed out that the theory had obvious important applications in biochemistry and physiology.

The Donnan principle was first used in the case of a protein by Procter[45] and by Procter and Wilson[46] in their investigations on the swelling of gelatin in acid. It was also used by Sörensen[47] in connection with his work on the osmotic pressure of egg-albumin solutions.

The most complete proof of the correctness of Donnan's theory of membrane equilibria has been furnished by Loeb,[48] who has shown

<center>TABLE XXVIII</center>

<center>CUBIC CENTIMETERS OF 0.1 N ACID IN COMBINATION WITH 1 GRAM OF ORIGINALLY ISO-
ELECTRIC CRYSTALLINE EGG ALBUMIN IN 100 C.C. OF SOLUTION (LOEB)</center>

pH	HCl	H_2SO_4	Oxalic Acid	H_3PO_4
4.2	1.15	1.15	1.18	3.8
4.0	1.7	1.7	2.6	5.3
3.8	2.3	2.3	3.7	6.8
3.6	2.9	2.9	5.0	8.6
3.4	3 5	3.5	6.3	10.6
3.2	4.2	4.3	8.0	13.1
3.0	5.0	5.1	9.5	16.1
2.8	5.8	5.9	11.1	19.3
2.6	6.7	6.5	13.3	22.9
2.4	7.6	7.0	16.0	

that the theory is undoubtedly applicable to solutions of the proteins gelatin, egg albumin and casein; and further work performed by Hitchcock[49] in Loeb's laboratory has extended the application to

[45] J. Chem. Soc., **105**, 313, 1914.

[46] J. Chem. Soc., **109**, 307, 1916.

[47] Z. physiol. Chem., **106**, 1, 1916; Compt. rend. trav. Lab. Carlsberg, **12**, 262, 1917 [in English].

[48] Proteins and the Theory of Colloidal Behavior. McGraw-Hill Book Co., New York, 1922.

[49] J. Gen. Physiol., **4**, 597, 1921–22; *ibid.*, **6**, 35, 1922–23.

edestin and serum globulin. From their investigations these authors conclude that proteins behave like crystalloids both in regard to their chemical reactions and their solubility, but on account of the large size of their ions they easily fulfill the condition for the establishment of a Donnan equilibrium, namely, that the protein ion is prevented from diffusing through membranes or gels which are easily permeable to the smaller crystalloidal ions.

TABLE XXIX

Cubic Centimeters of 0.1 N Acid in Combination with 1 Gram of Originally Iso-Electric Gelatin in 100 c.c. of Solution (Loeb)

pH	HCl	H_2SO_4	Oxalic Acid	H_3PO_4
4.0	....	2.7	3.9	6.95
3.8	3.9	3.75	5.5	9.4
3.6	4.8	4.8	7.3	12.3
3.4	5.6	5.75	9.1	15.2
3.2	6.4	6.75	11.0	18.0
3.0	7.2	7.5	13.15	20.7
2.8	7.9	8.25	15.3	23.6
2.6	8.35	8.8	17.1	26.2
2.4	8.5	9.3	18.0	

We are now ready to consider the query cited above: Why are proteins termed colloids? According to Loeb, the colloidal behavior of proteins is manifested only in those properties which depend on the presence of a non-diffusible ion; namely, membrane potentials, osmotic pressure, swelling, and that form of viscosity which is due to the swelling of submicroscopic particles. In other words, Loeb and others have shown that the colloidal behavior of proteins can be explained quantitatively on the basis of Donnan's theory of membrane equilibria.

The Donnan theory will now be presented in some detail, together with its application to the study of protein systems.

56. Donnan's Membrane Equilibrium.—As noted above, Donnan, in 1911, developed a theory to account for the type of equilibrium which results when a membrane separates two electrolyte solutions one of which contains an ion unable to diffuse across the membrane, while all the other ions can diffuse with ease, the result being an unequal distribution of the diffusible ions on the opposite side of the mem-

brane. As an example of this phenomenon, Donnan uses an aqueous solution of a sodium salt, NaR, such as Congo red, in contact with a membrane which is impermeable to the anion R^- and the non-ionized salt NaR, but will allow water and Na^+, or any other ion, to pass freely through it. The membrane separates the Congo red solution from an aqueous solution of sodium chloride, which will diffuse from its Solution II into the Solution I of NaR.

In view of the far-reaching effects of Donnan's membrane equilibria theory, it may be best to quote his own words:

"We suppose that the membrane (indicated in the following diagram by a vertical line) is impermeable for the anion R of a sodium salt NaR (and also for the non-dissociated part of the salt NaR), but permeable for all the other ions and salts to be considered in this connection. . . .

"Suppose that in the beginning we have a solution of NaR on one side of the membrane (indicated by the vertical line) and of NaCl on the other side:

$$\begin{array}{c|c} Na^+ & Na^+ \\ R^- & Cl^- \\ (1) & (2) \end{array}$$
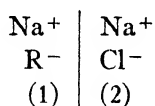
In this case NaCl will diffuse from (2) to (1). In the end the following equilibrium will result:

$$\begin{array}{c|c} Na^+ & Na^+ \\ R^- & \\ Cl^- & Cl^- \\ (1) & (2) \end{array}$$
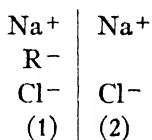
"When this equilibrium is established the energy required to transport reversibly and isothermally 1 gram molecule Na^+ from (2) to (1) equals the energy which will be gained by the corresponding reversible and isothermal transport of a gram molecule Cl^-. In other words, we consider the following infinitely small and reversible change of the system:

$$\left\{ \begin{array}{l} dn \text{ Mol } Na^+ \ (2) \rightarrow (1) \\ dn \text{ Mol } Cl^- \ (2) \rightarrow (1) \end{array} \right\}$$
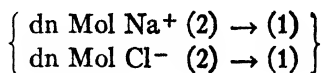
"The energy which can be gained in this way (i.e., the diminution of free energy) is zero, hence:

$$dn \cdot RT \log \frac{(Na^+)_2}{(Na^+)_1} = dn \cdot RT \log \frac{(Cl^-)_2}{(Cl^-)_1} = 0,$$

or

$$(Na^+)_2 \cdot (Cl^-)_2 = (Na^+)_1 \cdot (Cl^-)_1$$

where the brackets signify molar concentrations."

This last equation is the equilibrium equation, which signifies that the product of concentrations of a pair of diffusible ions (in the case Na^+ and Cl^-) on one side of the membrane will be the same as the product of the concentrations of the same ions on the other side of the membrane. But since the concentration of the cation Na^+ in (1) must be equal to the sum of the anions present $(R^- + Cl^-)$ in order to maintain electrical equilibrium, whereas on the other side (in [2]) the concentration of Na^+ is only equal to (Cl^-), it follows that

$$(Na^+)_1 > (Na^+)_2 \quad \text{and} \quad (Cl^-)_1 < (Cl^-)_2.$$

By the use of the above theory it is possible to account for the existence of different concentrations of diffusible ions on the two sides of a semi-permeable membrane, for example, within and without the red cells of the blood. This difference in concentration of diffusible ions leads to a *potential difference* (electrical pressure measure in volts) on opposite sides of the membrane, and Donnan has shown that this difference must be (on the basis of Nernst's well-known formula)

$$\pi_1 - \pi_2 \frac{RT}{F} \log \frac{(Na^+)_2}{(Na^+)_1} = RT \log \frac{(Cl^-)_1}{(Cl^-)_2}$$

or, since $\dfrac{RT}{F} = 58$ millivolts (at room temperature), the potential difference on the opposite sides of the membrane is equal in millivolts:

$$\pi_1 - \pi_2 = 58 \log \frac{(Na^+)_2}{(Na^+)_1} = 58 \log \frac{(Cl^-)_1}{(Cl^-)_2}.$$

Since many compounds in the body are of the form NaR, e.g., sodium proteinate, and the living tissues abound in semi-permeable membranes, it follows that Donnan's theory of membrane equilibria a wide application in biochemistry and biophysics.

In 1914 Proctor[50] proposed a theory of swelling based on Donnan's theory. He assumed a gelatin sol to contain gelatin ions which cannot diffuse into the surrounding water through the surface of the piece of gelatin. Thus the surface is equivalent to a membrane impermeable to gelatin ions, but permeable to other ions. Hence, Donnan's membrane equilibrium laws should obtain in this case, which implies that the total concentration of all ions within the gelatin is different from that outside the gelatin, and consequently, the osmotic pressure is different, and this difference of osmotic pressure causes the gelatin to attract water and therefore to swell. This theory has been verified by several investigators, notably Jacques Loeb. In fact, it was this theory of Proctor's on the swelling of gelatin that led Loeb to undertake his great work on proteins and their colloidal behavior. Loeb studied the equilibria resulting from the separation of solutions of protein salts from protein-free aqueous solutions by collodion membranes, and from the results that he obtained, he became convinced that the colloidal properties of proteins (membrane potentials, osmotic pressure, swelling, and that form of viscosity which is due to the swelling of submicroscopic particles) can be explained on the basis of Donnan's theory of membrane equilibria.

57. Proteins and Membrane Potentials.—In order to correlate Loeb's experiments with the Donnan theory, let us suppose hat we have two systems, (A) gelatin chloride in equilibrium with HCl, and (B) gelatin sulphate in equilibrium with H_2SO_4, in both of which the protein salt is separated from the acid by a collodion membrane, which is permeable to all the ions save that of the gelatin. The two systems may be expressed mathematically as follows:

$$
\begin{array}{c|c}
\begin{array}{c} z\,G^+ \\ y\,H^+ \\ y+z\,Cl^- \\ (1) \end{array} & \begin{array}{c} x\,H^+ \\ x\,Cl^- \\ (2) \end{array}
\end{array}
\qquad\qquad
\begin{array}{c|c}
\begin{array}{c} z\,G^+ \\ y\,H^+ \\ \dfrac{y+z}{2}SO_4{}^- \\ (1) \end{array} & \begin{array}{c} x\,H^+ \\ \dfrac{x}{2}SO_4{}^- \\ (2) \end{array}
\end{array}
$$
$$\quad\quad\quad\quad A \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad B$$
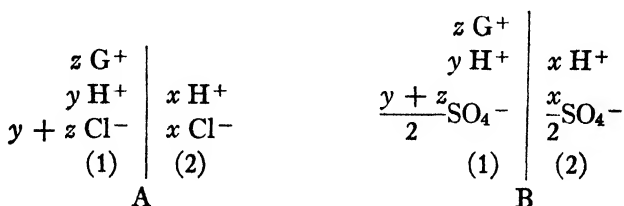
In which x, y, and z denote the various ionic concentrations. It is apparent from the reasoning given in connection with the Donnan theory that at equilibrium, in system (A), $x^2 = y(y + z)$, and in sys-

tem (B), $x^3 = y^2(y + z)$, if· we assume complete dissociation of the gelatin salts and the acids. It is evident, then, that on a basis of pure mathematical reasoning one can show that the valency of the diffusible ions will affect the final result. An unequal distribution of diffusible ions at equilibrium on two sides of a membrane of the type indicated must result in a difference in electrical potential between the two solutions separated by the membrane. Now, according to the Donnan theory, this potential difference is represented quantitatively by the expression $\frac{RT}{F} \cdot \log_e \frac{x}{y}$, which in the case of the system (A) becomes

$$\frac{RT}{F} \cdot \log_e \frac{\sqrt{y(y + z)}}{y} = \frac{RT}{2\,F} \cdot \log_e \left(1 + \frac{z}{y}\right),$$

and in the case of system (B)

$$\frac{RT}{F} \cdot \log_e \sqrt[3]{\frac{y^2(y + z)}{y}} = \frac{RT}{3\,F} \log_e \left(1 + \frac{z}{y}\right).$$

In other words, the ratio of the values of the potential differences in the two systems is 3 : 2.

By a very ingenious arrangement, Loeb succeeded in measuring potentials deducible mathematically as given above. The apparatus used is shown in Fig. 34. It consists essentially of a pair of saturated calomel electrodes, each having a capillary arm filled with a saturated solution of KCl. The end of one capillary arm was dipped into the gelatin solution in the collodion bag, and the end of other into the external solution. The calomel cells were then connected to a Compton electrometer and the potential difference measured.

After equilibrium had been attained (about 18 hours at 24° C.) the membrane potential was obtained by measuring the E.M.F. of the cell:

Calomel electrode	Saturated KCl	Gel. acid solution	Collodion membrane	Outside acid solution	Saturated KCl	Calomel electrode

On plotting the observed membrane potentials in millivolts against the pH of the acid protein solution, it was found that the values given by acids with monovalent ions, such as HCl, HBr, HI, HNO_3, acetic, propionic, and lactic acids, all lie on one curve, whereas the potentials with acids with divalent ions, such as H_2SO_4 and sulphosalicylic acid,

lie on a common but different curve. In all cases the curves show a minimal membrane potential at the iso-electric point of the protein, and all rise rapidly with diminishing pH until a maximum is reached at
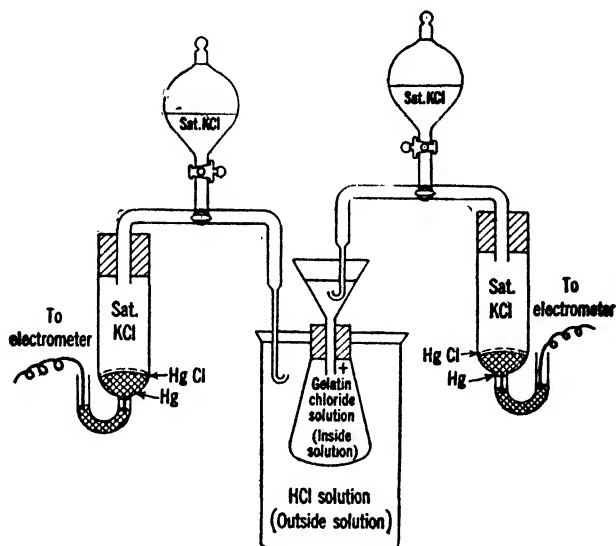
FIG. 34.—Method of Measuring the Potential Difference between a Gelatin Salt Solution in a Collodion Bag and the Outside Acid (having the same anion) Solution in a Beaker. (After Loeb.)

approximately pH 4.0, and gradually fall again with a further diminution of the pH. The following table (Fig. 35) illustrates these points:

It has already been shown on mathematical grounds that at the same pH of the protein solution the ratio of the potential difference of a divalent ion over that of monovalent ion must be as 2 : 3 or 0.66. Table XXX shows how nearly the experimental results accord with this theoretical requirement and leave little doubt in regard to the correctness of the Donnan theory of membrane equilibria in connection with the origin of the potential difference between protein solutions and outside aqueous solutions:

We shall now briefly consider the effect of adding neutral salts to a protein-acid system. We have already learned (p. 289) that the addition of neutral salts to solutions of protein salts depresses the osmotic pressure and viscosity of these solutions, and that the addition of neutral salts to a gel depresses the swelling of the latter (except

when the solutions and gels are at the iso-electric point). Accordingly Loeb undertook to find out whether or not the addition of a salt to a protein solution depresses also the potential difference across a membrane, and whether this was also due to the depression of the value of the pH inside minus the pH outside. The correctness of this theory of the action of neutral salts was also amply proved by Loeb. In brief, he was able to show by experiment that neutral salts depress the potential difference which exists at the point of equilibrium between a protein acid (chloride) solution contained in a bag and an outside

TABLE XXX

MEMBRANE POTENTIALS FOR DIBASIC AND MONOBASIC ACIDS (AFTER LOEB)

pH	Dibasic acids, millivolts	Monobasic acids, millivolts	Ratio $\dfrac{\text{dibasic}}{\text{monobasic}}$
2.4	7.6	11.4	0.67
2.6	9.6	14.8	0.65
2.8	11.6	18.0	0.64
3.0	13.6	21.6	0.65
3.2	15.8	24.8	0.64
3.4	18.0	28.0	0.62
3.6	19.8	31.0	0.64
3.8	21.2	34.2	0.62
4.0	21.6	35.5	0.61
4.2	20.8	34.8	0.60
4.4	19.2	31.0	0.62

aqueous solution (without protein). This depression in the potentital difference can be calculated by means of Nernst's formula[51] on the assumption that the potential difference is due to a difference in the hydrogen-ion concentration on the opposite sides of the membrane. This difference in hydrogen-ion concentration is due to Donnan's equilibrium.

58. Osmotic Pressure of Protein Solutions.—The same experiments that were used for the measurement of membrane potentials were also used for measuring the osmotic pressure of proteins. Most of the results were obtained with gelatin chloride and various acids. These are shown in Fig. 36. The ordinates are the observed osmotic

[51] J. Gen. Physiol., **3**, 667, 1921.

pressures in terms of millimeters of a column of water, and the abscissæ are the pH of the gelatin solution at equilibrium. Here, again, we observe the phenomenon of a minimum point in a physical property (osmotic pressure) of a protein corresponding with the iso-electric
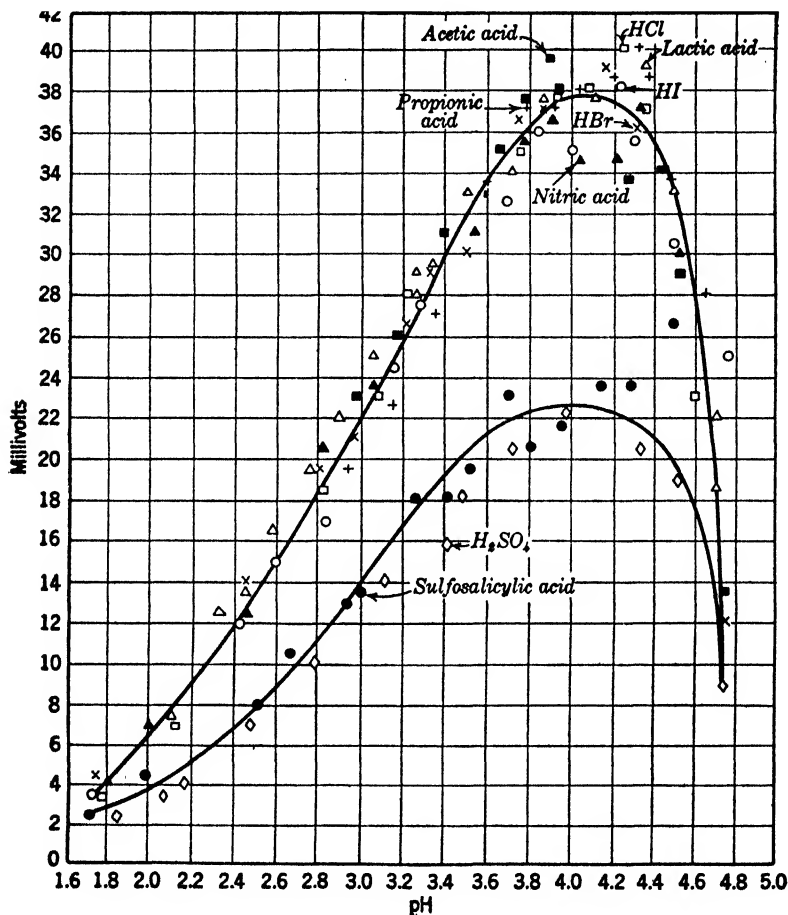
FIG. 35.—Showing that Only the Valency of the Anion of an Acid Influences the Membrane Potentials of Protein Solutions.   (After Loeb.)

point. Note that it rises upon the addition of acid until a maximum is reached at pH 3.3, and that, upon the further addition of acid, the osmotic pressure diminishes again. It is also noticeable that the values for all the monobasic acids lie practically on one curve. The osmotic pressure curves for the two strong dibasic acids, H_2SO_4 and sulpho-

salicylic acid, also fall on one curve, which is, however, only about half as high as the curve for the monobasic acids for the same pH.[52]

Figure 37 illustrates the influence of weak dibasic and tribasic acids on the osmotic pressure of proteins. Note that the descending branches
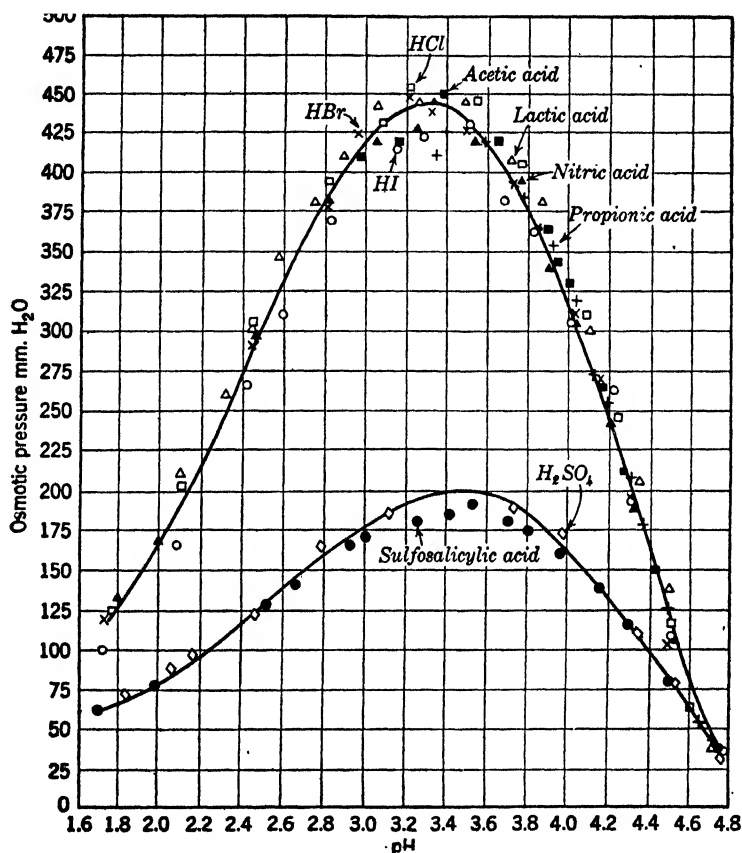
FIG. 36.—Proof of the Valency Rule for the Influence of Acids on the Osmotic Pressure of Gelatin Solutions. (After Loeb.)

of the curves for the three acids used (succinic, tartaric, and citric) are identical with the corresponding part of the curve for HCl for pH below 3.0, while above this point the curves are slightly lower in the order of their relative strengths. (See discussion under Article 54, p. 290).

It is interesting to note that the above figures are so exact that it

[52] Loeb and Kunitz, J. Gen. Physiol., 5, 665, 1922–23.

has been found that it would be possible to determine the relative strengths of dibasic or tribasic acids from their influence on the osmotic pressure of standard protein solutions.

The data given in Figs. 36 and 37 show that the valency, but not the chemical nature of the anion of the acid, influences the osmotic
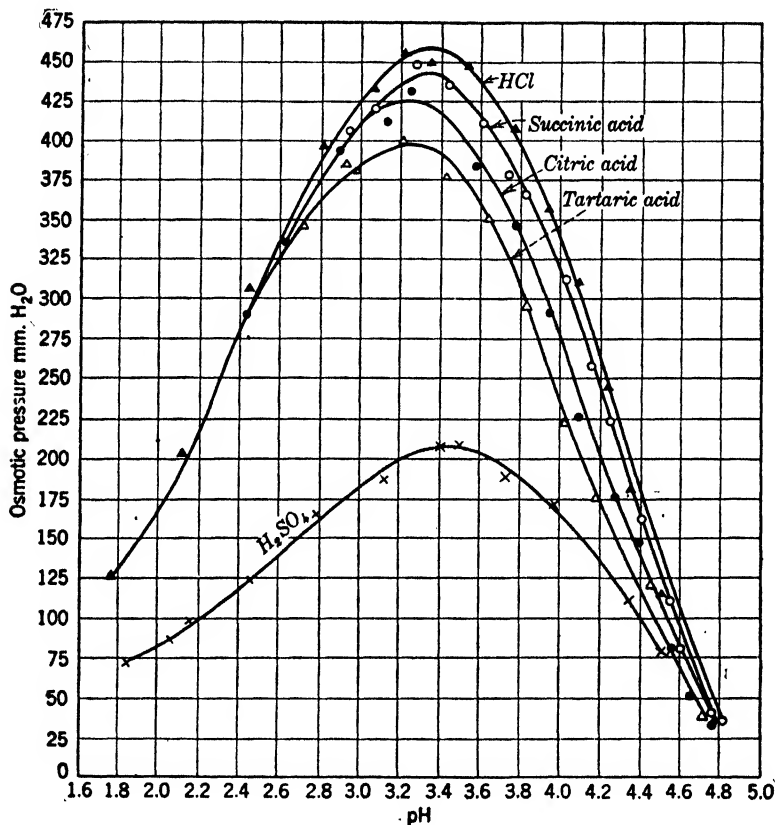
FIG. 37.—Influence of Weak Dibasic and Tribasic Acids on the Osmotic Pressure of Protein Solutions.   (After Loeb.)

pressure, which is to be expected if this influence of the acid is due to the Donnan equilibrium.

It is important to perceive that these curves of observed osmotic pressure are not exclusively the expression of the osmotic pressure due to the protein particles, protein molecules, and protein ions alone, but

are also the result of the demonstrable unequal concentrations of the crystalloidal ions on the opposite sides of the membrane, caused by the establishment of a Donnan equilibrium. In other words, the observed osmotic pressure of a protein solution needs a correction due to the Donnan equilibrium before it is possible to speculate on the cause of the influence of acids on osmotic pressure. Loeb[53] has calculated the value of this correction, and has shown that when the correction is applied the influence of the pH of the acid on the osmotic pressure is entirely, or practically entirely, due to the excess of the concentration of crystalloidal ions inside the membrane over that outside. For the numerical value of this correction the reader is referred to the above-mentioned paper, also the chapter on proteins in Bogue's "Colloidal Behavior," published by the McGraw-Hill Book Co., New York.

Loeb summarizes his researches on this phase of the colloidal behavior of proteins in the following terms: " The so-called colloidal behavior of proteins, as far as membrane potentials and osmotic pressure are concerned, is merely the result of an equilibrium condition of classical chemistry, which results in an excess of the concentration of crystalloidal ions inside the protein solution over that of an outside aqueous solution, when the two solutions are separated by a membrane that is permeable to crystalloidal ions but impermeable to protein ions. The colloidal behavior of proteins depends, therefore, entirely on the relative non-diffusibility of protein ions through membranes that are easily permeable to crystalloidal ions."

59. Nature and Structure of Gels.—One of the best-known properties of emulsoids is that of gelatinization, or, gel-formation. In other words, the one property that most sharply distinguishes sols from true solutions is their ability to "set" into a gelatinous semi-solid, mass, known as a *gel*, without any change in chemical composition, or proportions, of the component parts of the system. Graham applied the term "gel" to all the products of transformations of sols. Since his time, some scientists have used the term gel to signify the firm mass which one obtains when a substance like gelatin is dissolved in warm water and allowed to cool. It is better, however, to use the term "jelly" for this latter type, in order to avoid confusion.

There is also a possibility of one's confusing the term "gelatinous precipitate" with the terms gel and jelly. When certain colloidal

[53] Science, **56**, 731, 1922.

solutions are caused to coagulate under suitable conditions we may get a gelatinous precipitate with a supernatant liquid, which is usually a dilute solution of the same material as the precipitate. It is evident, then, that the terms "gel," "jelly," and "gelatinous precipitate" must not be confounded.

Concerning the structure of gels a great many investigators have occupied themselves without coming to any very definite understanding. Biologists, as well as chemists and physicists, have been engaged in this work. According to the majority of these investigators, however, a gel is a two-phase system in which the dispersing agent is a solid, and the disperse phase a liquid. Let us consider the case of gelatin jelly, which is obtained on cooling a fairly concentrated solution of gelatin. It has the properties of a solid, in that it holds its shape and resists shearing stresses; it possesses considerable elasticity. It may be regarded as formed by the separation of the cooling solution of gelatin into two phases: a solid phase, rich in gelatin, which forms a meshwork in the interstices of which is imbedded a very much more liquid second phase. The process of gelation has been studied with the ultramicroscope, and it has been found that the structures of various gels may be quite different in details. For example, by observing solutions of pure soap, it has been noted that as the solution cools the gel at first forms a polarized cone of light, but the very fine particles which are responsible for the effect soon increase in number and size so that they obstruct one another in their Brownian movement and adhere, giving the appearance of fine felt-like threads throughout the mixture (Bradford). Other gels are formed of very thin, needle-like crystals which penetrate the mass in all directions, and which hold the more dilute solution entangled between them. The clotting of fibrinogen is an important example of this type of gel formation. In this it differs from the clotting of casein, which is a typical jelly of the non-crystalline form. If rennin is added to a solution of casein under suitable conditions, the casein is converted into an insoluble form, paracasein, which gradually forms a jelly. If this type of gelation is watched under the ultramicroscope, it will be observed that before the gel formation starts the field is homogeneous and dark, but as soon as the process begins there appears a diffuse, very faint light in the field, but no visible, distinct points; as the gel formation goes on, distinct points in active Brownian movement appear in the field; as these points grow larger and larger, their movement slows

down, until finally the gel is observed to consist of an immense number of small clumps distributed throughout the entire mass and having no Brownian movement. In this case, which is typical of most gels, the gelation appears to be due to the formation of an insoluble precipitate, which, instead of flocking out of the solution, forms a sort of impervious sponge work, the more fluid phase being enclosed in the meshes. This liquid phase will necessarily consist of a saturated solution of the substance, in this case paracasein, which has made the precipitate. Other substances will also be held in solution in the liquid pockets.

Gels may be either rigid, as in the case of silicic acid, or elastic, as are those of gelatin, agar, egg albumin, etc. The latter is the more common type of gels among organic colloids. They can be easily changed in shape or form, without any change in the total volume.

Gels may also be classified as reversible and irreversible. The former are reconvertible into a sol state by warming or other treatment, and form a jelly again on cooling. Gelatin and agar are the best-known examples of reversible gels. The irreversible gels are not reconvertible into sols. Blood, when it clots, forms an irreversible gel.

60. Imbibition and Swelling of Gels.—Practically all substances that form elastic gels exhibit the property of swelling in a suitable liquid. Thus, dry gelatin, fibrin, and starch will swell in water at ordinary temperatures, forming jellies that are peptized at higher temperatures, forming sols. Most proteins will swell in water, but not in the common organic solvents—alcohol, ether, chloroform, benzene, toluol, xylol, or turpentine. Vulcanized rubber swells in the organic solvents, but not in water.

On swelling, gels imbibe considerable quantities of the liquid, and the "cells" become distended so that the mass of the material expands very considerably. Perhaps the most remarkable feature of the phenomenon is the great pressure exerted in the process of swelling, or conversely, required to express the liquid after it has been imbibed. *Laminaria* (a genus of sea weeds), under a pressure of 42 atmospheres will still expand 16 per cent, while, with a pressure of 1 atmosphere the expansion amounts to 330 per cent.

During the swelling of gels by imbibition, the total volume—gel plus fluid—becomes less, although the volume of the gel itself increases enormously. For example, in a given mixture of gelatin and water, it was found that after the gelatin had ceased to swell, the jelly occupied 2 per cent less space than the total volume of the original gelatin and

water. Now, in order to compress water to the extent implied in this change in volume, a pressure of some 400 atmospheres is necessary, so that it is plain that considerable heat must be involved in the process of imbibition. On the other hand, if a swollen gel is placed in a dry atmosphere, it rapidly loses water, with a corresponding decrease in volume. A large amount of the water is held quite loosely, so that the evaporation from the gel proceeds fairly rapidly at first, and then more and more slowly, the last traces of water being held by the gel with extreme tenacity. On this account some gels must be heated for many hours at a temperature of 120° C. in order to be thoroughly dried.

It is interesting to note that, in the process of swelling, considerable heat is set free and considerable quantities of energy suffer conversion. The swelling of seeds, for example, can lift great weights. The Egyptians used the phenomenon of swelling for quarrying purposes when they drove wooden wedges into rocks and then caused the wedges to swell by pouring water upon them.

In all cases of swelling of emulsoids, the swollen body possesses much less cohesion (internal attraction of the molecules) and greater ductility than it had before swelling. In the case of fluid absorption by non-swelling substances, the cohesion is too great to permit the molecules to be forced apart by the fluid. In the case of colloidal swelling in water, the particles are forced apart just enough to make the tissue soft and elastic. If the separation of the particles were to continue until the molecules were segregated, a true solution would be produced. Some colloids form a jelly at first, and then gradually a true solution; for example, gum arabic on being mixed with water swells into a stiff jelly, then into a soft jelly, and finally completely dissolves, forming a true solution.

The degree of swelling which a gelatin plate undergoes in water is greatly enhanced by the addition of small amounts of acid or alkali to the water, the minimal imbibition of water being at the iso-electric point—pH 4.7.

Procter and Wilson[54] have shown that the influence of an acid on the swelling of gelatin is purely as osmotic effect. "The acid, combining with the gelatin, causes salt formation, the gelatin ions being prevented from diffusing by the cohesive forces between the gelatin ions or mol-

[54] J. Chem. Soc., 109, 307, 1916.

ecules of the gel. Since the gel is freely permeable to water and crystalloidal ions, such as H^+ and Cl^-, the non-diffusibility of the gelatin ions causes the establishment of a Donnan equilibrium between gelatin and outside solution, as a result of which the total molar concentration of all the diffusible crystalloidal ions is greater inside than outside the gel. This causes the influence of the acid on the swelling of gelatin and this influence is the same as that on osmotic pressure, for the reason that the influence of acid on swelling is also an osmotic-pressure effect. The difference between the effect of acid on the osmotic pressure of gelatin solutions (in collodion bags) and on the swelling of gelatin gels is simply this—that in the former case the diffusion of the gelatin ions is blocked by the collodion membrane, and in the latter case by the cohesive forces between the gelatin molecules or gelatin ions of the gel."

Since cohesive forces are also a limiting factor in the swelling of gels, it is evident that the phenomenon is rather complicated. Gels will imbibe a certain amount of water at the iso-electric point. At this point, the gels are very slightly ionized; hence, it is evident that the swelling which takes place in electrically neutral gels cannot be attributed to the Donnan equilibrium. It must, therefore, be due to forces of attraction between certain groups of the molecules of the gel and water. The additional swelling induced by the addition of acid is, however, as Procter and Wilson have shown, an osmotic phenomenon due to the excess in the concentration of the ions of the acid inside over that outside, which effects a diffusion of water into the gel. According to Loeb, "the hydrostatic pressure of the water will force the molecules apart and this will cause an increase in the forces of cohesion, which will oppose further swelling."

The effect of a slight increase in pH on swelling is very pronounced. One gram of powdered gelatin at the iso-electric point (4.7) will imbibe 7 grams of water, whereas at pH 3.2 it will imbibe 35 grams of water. At higher pH values than 3.2, however, there is a gradual decrease in the amount of water imbibed.

The idea that swelling was a consequence of the Donnan equilibrium was advanced by Procter and Wilson in 1916 and was later confirmed by Loeb,[55] who showed that the Hofmeister ion series, which had been obtained previously in experiments on swelling of proteins,

[55] Proteins, Chapter IX.

has no real significance in dilute solutions if the effect of ions is compared at the same pH. His method of calculating the effect of an acid on the swelling of a gel on the basis of Donnan's equilibrium theory was similar to that which he used for calculating the osmotic pressure, but it was further complicated by the necessity of introducing the cohesive forces of the jelly. For the details of the calculation the reader is advised to consult Loeb's book, "Proteins and the Theory of Colloidal Behavior," page 190. Loeb reasoned that, in case the influence of an acid on swelling is due to the Donnan equilibrium, the influence of different acids must depend solely on the valency and not on the nature of the anion of the acid. The results he obtained prove that his assumption was correct, since the curves obtained for the swelling of gelatin in different acids were similar to the osmotic-pressure curves given in Figs. 36 and 37. In a few instances abnormal results were obtained in the case of swelling, that were not obtained in either membrane potentials or osmotic pressure. For example, higher results were obtained with acetic acid than were obtained with such monobasic acids as HCl or HNO_3. Loeb suggests that this excessive effect of acetic acid on swelling is due to a diminution of the cohesion of the gel, caused by the high concentration of this acid required to bring the pH to 3.2.

It has likewise been shown that the swelling of gels is depressed by salts in the same way as is the osmotic pressure, which is another proof of the correctness of the theory, since both phenomena must depend on the excess concentration of diffusible ions inside the membrane or jelly. It has also been proved that the valence effect demanded by the theory holds on the alkaline side of the iso-electric point, and that no Hofmeister series of cations was exhibited.

In short, then, the results that have been obtained up to date are sufficient to show that the swelling of gels in acids and alkalies, and their depression by salts, can be most satisfactorily explained as a result of the Donnan equilibrium.

Numerous other theories[56] have been advanced to explain the phenomenon of swelling, but none is entirely satisfactory.

T. Brailsford Robertson[57] has advanced an interesting interpreta-

---

[56] These theories have been summarized, and their limitations pointed out, in a paper by Bartell and Sims: J. Am. Chem. Soc., 44, 289, 1922.

[57] Professor of Physiology and Biochemistry at the University of Adelaide, South Australia.

tion of the influence of acids and alkalies on the swelling of proteins. He postulates: "The increased *swelling-capacity* of gelatin in solutions of acids or alkalies is merely the expression of the fact that the ionization of the protein salt leads to an increase in the number of colloidal particles per unit volume of the jelly, and possibly, also, in part, to the fact that protein ions have a greater affinity for water than undissociated protein molecules."

61. Viscosity of Protein Solutions.—We are also indebted to Jacques Loeb for the discovery of the fact that the influence of hydrogen-ion concentration, valence of ions, and concentration of added salt on the viscosity of protein solutions is remarkably similar to the influence of these factors on membrane potentials, osmotic pressure, and swelling. The individual proteins, however, were found to vary considerably in this property. For instance, the viscosity of purified egg albumin was found to be much lower than that of gelatin, and changes in pH had but little effect on its viscosity as compared with that of gelatin. According to Loeb, the high viscosity of gelatin solutions is due to the presence of submicroscopic particles of gelatin which occlude water. Since the amount of water occluded by solid jelly is regulated by the Donnan equilibrium, this will explain why the influence of electrolytes on the viscosity of a gelatin solution is similar to that of electrolytes on the viscosity of *suspensions* of powdered gelatin in water. Experiments on suspensions of other protein salts also support this idea. From these data, Loeb draws the conclusion that there are two sources of the viscosity of protein solutions; one being the free ions and molecules of the protein; the other, the submicroscopic solid particles in the solution. The viscosity caused by the ions and molecules is the general viscosity, which is common to all solutions; it is of a low order of magnitude like that of crystalloidal solutions. The high viscosity due to the submicroscopic particles of protein is specifically colloidal. General viscosity prevails in dilute solutions of crystalline egg albumin and metal caseinates (where the metal is monovalent) under ordinary conditions of hydrogen-ion concentration and temperature. Under the same conditions the specific colloidal viscosity prevails in gelatin solutions, and in solutions of acid salts of casein, also in solutions of crystalline egg albumin at pH below 1.0 and higher temperatures.

The above data and many similar observations, which cannot be included because of lack of space, made it possible for Loeb to prove

that the similarity between the influence of electrolytes on the viscosity of gels and the influence of electrolytes on osmotic pressure is due to the fact that the influence on viscosity in such cases is in reality an influence on the swelling of submicroscopic protein particles, i.e., a function of osmotic pressure.

62. Syneresis.—If a gel is left undisturbed for a certain period of time, protected not only against infection with microörganisms, but also against evaporation, a separation into two phases takes place. Clotted milk separates into whey and curd; blood clot separates into serum and a relatively solid mass; agar slants exude fluid droplets. In fact, this segregation of liquid occurs in all elastic gels, and even in some rigid gels. The liquid is often called condensation water. This, however, is a designation that is likely to mislead one, for the liquid given off from aqueous gels is not merely water, but a solution of all the constituents of the gel—colloidal and crystalloidal. The relative proportions are, of course, quite different. This "serum" secretion by colloids was observed by Graham, who called it "syneresis."

The phenomenon is common to all kinds of colloidal gels, but the concentration at which it is most marked differs considerably. The rigid gels, like silicic acid, show marked syneresis at high concentrations, whereas with elastic gels, like those of gelatin and agar, it is best observed at greater dilutions.

This property of gels, in conjunction with that of osmotic pressure and swelling, still offers a large and promising field for research, especially in relation to its bearing on such problems as those of glandular secretion.

63. Summary of the Colloidal and Crystalloidal Behavior of Proteins.—The experimental data which have been presented in the foregoing pages show that the proteins possess both colloidal and crystalloidal properties.

Evidence has been presented to show that the four properties— membrane potentials, osmotic pressure, swelling, and that type of viscosity which depends on the swelling of submicroscopic solid particles in the solutions—are the specific properties which constitute the colloidal behavior of proteins; and this behavior has been explained quite satisfactorily on the basis of amphoteric ionization and the Donnan equilibrium.

Electrolytes have a similar influence on all of the four properties mentioned, and it may be summarized as follows:

(1) The addition of acid or alkali to iso-electric protein increases at first the values of these four properties until a maximum is reached, after which the addition of more acid or alkali diminishes their values again.

(2) This influence of acids and alkalies depends only on the valency, and not on the chemical nature of the ions, as was suggested by Hofmeister.

(3) When the anion of the acid or the cation of the alkali is divalent the values of these four properties are considerably less than when the ion is monovalent.

(4) The addition of a neutral salt to a protein solution or gel—not at the iso-electric point—depresses the value of the four properties, and this depressing effect increases with the valence of that ion of the salt which has the opposite electrical charge to that of the protein ion.

There are other properties of proteins in which their behavior is similar to that of crystalloids, namely, their solubility relations and their reactions with acids and bases.

The colloidal nature of the proteins is attributed to the large size of their ions, which enables them to fulfill the conditions necessary for the establishment of a Donnan equilibrium, namely, that the protein ion is prevented from diffusing through membranes or gels which are easily permeable to the smaller crystalloidal ions.

CHAPTER XII

CATALYSIS AND VELOCITY OF CHEMICAL REACTIONS

1. General Concept of Catalysis.—One of the most important factors in the process of chemical change is the speed with which it takes place, *the velocity of the reaction.* In inorganic chemistry, the question of the rate of a chemical change does not arise as often as it does in organic or physiological chemistry, because in general the reactions are so rapid that it is not possible to measure their speed. The neutralization of an acid by a base, for instance, as shown by the change of color of some indicator, is almost instantaneous. There are a few well-known reactions which are exceptions to this rule; among these we may cite the oxidation of sulphur dioxide, and the decomposition of hydrogen peroxide. In organic chemistry, on the other hand, slow chemical reactions are frequently encountered. Thus, the reaction between an alcohol and an organic acid, forming an ester and water, proceeds very slowly under ordinary conditions, and one has no trouble in following the progress of the reaction. It is interesting to note, however, that chemical changes that are accomplished with difficulty outside the living cell are often accomplished inside it with apparent ease. This superior rate of reactions within a living cell is due to the fact that these reactions are enhanced or *catalyzed* with the assistance of enzymes, or catalysts, of cellular origin.

The characteristic action of a catalyst is to affect the rate of a reaction. It is a well-known fact that a chemical change which of itself proceeds with extreme slowness may be greatly accelerated through the presence of some apparently foreign substance, the other conditions being unchanged. Often only a trace of the foreign material is sufficient to effect the transformation of large quantities of the reacting substances. For example, 1 gram of colloidal platinum will bring about the combination of 25,000 liters of hydrogen and oxygen without the activity of the catalyst being impaired. The fact that only the merest trace of a catalyst is necessary (in this case on 1 part in 25 million) in

314

a given process constitutes one of the important characteristics of catalytic reactions.

The phenomenon of catalysis has been known for a long time. It was defined by Berzelius as far back as 1835, following an extensive study on alcoholic fermentation. The words used by Berzelius are worth quoting: "We had made acquaintance with the fact that, for example, the change of sugar into carbonic acid and alcohol takes place in fermentation under the influence of an insoluble body, which we call 'ferment,' and also with the fact that this ferment could be replaced, although less effectively, by animal fibrin, coagulated plant albumin, cheese and similar substances, as well as with the experience that the process could not be explained by a chemical action between the sugar and the ferment analogous to double decomposition. Comparing it with known relations in the inorganic world, it was seen to be most like the decomposition of hydrogen peroxide under the influence of platinum, silver, or fibrin; it was, therefore, natural to suppose that the action of the ferment was an analogous one."

Berzelius defines the process as follows: "I will call it the *catalytic* power of substances, and the decomposition effected thereby, *catalysis*; just as we understand by *analysis* the separation of the constituents of substances by means of ordinary affinity. Catalytic power appears to consist essentially in the fact that substances are able to set into activity affinities which are dormant at this particular temperature, and this, not by their own affinity, but by their presence alone."

Turning to living nature, it is pointed out that "we have justifiable reasons to suppose that, in living plants and animals, thousands of catalytic processes take place between the tissues and the liquids in the formation of the great number of dissimilar compounds, for whose formation out of the common raw material, plant juice or blood, no probable cause could be assigned. The cause will perhaps in the future be discovered in the catalytic power of the organic tissues of which the organs of the living body consist."

The word catalysis is from the Greek *kata*, meaning "down," and *lysis*, "loosening." Literally it means unloosening, but it has gradually come to mean the hastening of a chemical reaction by a third substance, the catalyst, which, as has been stated, emerges at the end of the reaction practically unchanged in amount. In the terse language of Ostwald: "A catalytic agent is that material which affects the velocity of a chemical reaction without appearing in the final prod-

ucts." From this definition one might conceive the idea that the enhanced chemical reaction takes place in virtue of the mere presence of the catalyst; there is no doubt, however, that probably in all cases it has actually entered into the reaction at some stage of the procedure, and has become free again. In other words, it is probable that all catalytic agents are conductors of energy.

The phenomenon has become, in more recent times, of enormous importance in all branches of chemistry. In the field of biochemistry, catalysis is especially interesting, since, in all animal and vegetable organisms, there are going on unceasingly a multitude of catalyzed chemical reactions. Most of the chemical changes occurring in digestion and metabolism would take place far too slowly for the maintenance of normal life, if it were not for the fact that nature has provided for this contingency by having the living cells manufacture the necessary catalysts. These biological catalysts, or enzymes, will be discussed later.

In order that it may be easier to understand how catalytic agents may hasten reactions, we shall briefly consider the factors which determine the velocity of ordinary reactions.

2. Velocity of Chemical Reactions.—Chemical reactions may be roughly divided into two classes: (1) those termed "ionic," which are practically instantaneous reactions; and (2) those that take an appreciable time to attain equilibrium. It is evident that catalytic phenomena are concerned only with the latter, and particularly with that group of them which proceed so slowly that no reaction appears to be taking place at all.

The velocity of a chemical reaction is directly proportional to the chemical affinity, and inversely proportional to the chemical resistance, from which it follows that any alteration in either of these factors will affect the velocity of the reaction. For example, a rise of 10° C. temperature generally more than doubles the velocity of the reaction of a system, owing to the fact that heat not only increases chemical force but also decreases chemical resistance. However, it is evident that a rise in temperature is often most undesirable, especially in the case of reactions taking place in living organisms.

The basis of the study of the velocity of chemical reactions is furnished by the law of mass action, which we have already considered in connection with the conditions of equilibrium in the case of balanced reactions (Chapter IX). According to this law, the rate at which a

chemical reaction is taking place, at any given time, is proportional to the concentration of the reacting substances at that particular instant of time. Since the concentration of the reacting substances depends on the relation $\dfrac{\text{mass}}{\text{volume}}$, the law of mass action may be termed the law of concentration velocity.

3. Unimolecular Reactions.—The simplest type of chemical reaction is one in which only one substance undergoes change, and in which the velocity of the reverse reaction is negligible. This may be represented in simple form by the equation $A \rightarrow B$. We may represent the initial concentration of A by a mols. After a given time, t, a certain amount of B is formed, which may be represented by x. At this time the concentration of the unchanged A is, therefore, $a - x$.

The velocity of the reaction, or the rate at which B is being formed with time, may be expressed as dx/dt. From the mass law, this velocity at any given time is proportional to the amount of A present at that moment, and this will be equal to the amount of A at the start of the reaction minus the amount, x, transformed during the period, t, or

$$dx/dt = k(a - x) \tag{1}$$

The significance of the constant k (velocity constant) may be seen by transposing equation (1):

$$\frac{\dfrac{dx}{dt}}{a - x} = k. \tag{2}$$

In order to use equation (1) for any practical purpose, it must, of course, be integrated, so that the change during a measurable time can be investigated. In other words, the equation, which is merely the mathematical expression of the law of mass action, involves such infinitesimal quantities that it cannot be used directly in experimental work. With the aid of integral calculus, however, one can obtain from the equation an expression which involves only finite quantities, which can be measured. The simplest form of the integral is

$$k = \frac{1}{t} \log_e \frac{a}{a - x} \tag{3}$$

By changing this equation, which employs natural logarithms, to

one that employs logarithms to the base 10 (this is done by multiplying by 2.303), we get

$$k = \frac{2.303}{t} \cdot \log_{10}\frac{a}{a-x}. \tag{4}$$

In this expression, t, a, and x can be determined by experiment.

The hydrolysis of cane sugar is a typical unimolecular reaction.

4. Bimolecular Reactions.—When two substances react and both alter in concentration, the reaction is said to be bimolecular, or of the second order. Let a and b represent the initial molar concentrations of the two reacting substances, and let x denote the amount transformed in the interval of time $t\cdot$ then the velocity of the reaction may be expressed by the equation

$$v = \frac{dx}{dt} = k(a-x)(b-x) \tag{5}$$

The simplest case is that in which the substances are present in equivalent quantities. The velocity equation in this case becomes

$$\frac{dx}{dt} = k(a-x)^2 \tag{6}$$

which, on integration, gives the expression

$$\frac{1}{t}\frac{x}{a(a-x)} = k. \tag{7}$$

If the reacting substances are not present in equivalent amounts, the calculation is somewhat more complicated. In this case we must employ the equation

$$\frac{dx}{dt} = k \cdot (a-x)(b-x).$$

which on integrating becomes,

$$k = \frac{2.303}{t(a-b)} \cdot \log_{10}\frac{b(a-x)}{a(b-x)}. \tag{8}$$

This is the equation for a typical bimolecular reaction. The hydrolysis of an ester by an alkali is an example of this type of reaction.

Trimolecular reactions and reactions for still higher types may be

similarly formulated. For further details the student may consult standard texts on general physical chemistry.

5. The Hydrolysis of Cane Sugar: A Unimolecular Reaction.— The hydrolysis of cane sugar by acids can clearly be regarded as a unimolecular reaction, since it is completely accounted for, as regards its rate, by the change of concentration of the molecules of cane sugar. The change may be expressed by the equation

$$C_{12}H_{22}O_{11} + H_2O = 2C_6H_{12}O_6.$$
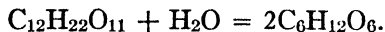
At first sight, this reaction appears to be a bimolecular one, since two molecules enter into it. However, the relative concentration of water is so large with respect to that of the cane sugar that during the reaction it undergoes very little change, the concentration of the cane sugar alone being converted. On applying formula (4) to this reaction, it is found that the constant k is practically the same at all stages of the reaction. The following table gives the results obtained on the hydrolysis of cane sugar with an acid, the varying concentrations of cane sugar being determined by means of a polarimeter, properly jacketed to maintain a constant temperature. Under these conditions the knowledge of the angle of rotation, determined at definite intervals, enables one to follow the course of the change and to evaluate k for each point.

INVERSION OF CANE SUGAR BY ACID

$a = 10,023$

t, in minutes	x	k
0	0.000	0.00000
30	1.001	0.00152
60	1.946	0.00156
90	2.770	0.00156
180	4.676	0.00151

6. Saponification of an Ester by an Alkali: A Bimolecular Reaction.— As a typical illustration of a bimolecular reaction, we may take the hydrolysis of an ester by an alkali, a reaction which has been thoroughly investigated by several chemists. When sodium hydroxide is added to a solution of ethyl acetate, for example, the ester is grad-

ually decomposed into sodium acetate and ethyl alcohol. The progress of the reaction is marked by the decreasing alkalinity of the solution; therefore, by taking a small measured portion from time to time, and titrating with standard acid, the velocity of the reaction can be quantitatively studied. The reaction, being a reversible one, may be written as follows:

$$CH_3COOC_2H_5 + NaOH \rightleftharpoons CH_3COONa + C_2H_5OH,$$
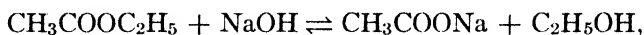
from which it is evident that in this case the concentrations of two substances undergo change during the reaction; hence the formula

$$\frac{dx}{dt} = k(a - x)(b - x)$$

applies here, where a and b are the initial quantities of ester and alkali, and x is the quantity of sodium acetate produced after an interval of time t from the start. To obtain the velocity of the reaction, we employ formula (8).

Arrhenius,[1] in certain experiments performed by him, used 0.02 M solutions of ester and alkali. These solutions were placed in flasks and warmed to 25° C. in a thermostat maintained constantly at that temperature; equal volumes were then mixed, and at frequent intervals a portion of the reaction mixture was removed and titrated rapidly with standard acid. A few of Arrhenius' results are recorded in the following table:

HYDROLYSIS OF ETHYL ACETATE BY SODIUM HYDROXIDE

t (minutes)	$a - x$	k
0	8.04	......
4	5.30	0.0160
6	4.58	0.0156
8	3.91	0.0164
10	3.51	0.0160
12	3.12	0.0162
		Mean = 0.0160

The numbers for $a - x$ in the second column represent the concentrations of NaOH (and of ethyl acetate) expressed as the number of cubic

[1] Zeit. phys. Chem., 1, 110, 1887.

centimeters of HCl required to neutralize 10 c.c. of the mixture. The reaction is very rapid, and therefore the experimental error is rather large; nevertheless, the values of k in the third column will be observed to differ very slightly from the mean value.

7. Velocity of Reaction and Temperature.—It is a matter of general experience that the rate of chemical reaction increases very greatly as the temperature rises. Thus, the rate of inversion of cane sugar is about five times as rapid at 50° C. as it is at 20° C. As a general rule, the velocity of reaction in *homogeneous* systems is doubled or trebled for an increase of 10° C.

It should be carefully noted that rise of temperature does not usually affect the form of the velocity equation; what is altered is the magnitude of the velocity constant. The accompanying table gives the value of k, the velocity coefficient, for the hydrolysis of cane sugar by acid at different temperatures.

Temperature	k
25°	9.67
40°	73.40
45°	139.00
50°	268.00
55°	491.00

According to the molecular theory, rise of temperature should increase the rate of chemical change through the accelerating of molecular movements; but this effect does not account for all of the facts, for if only the velocity of the molecular movement was concerned in the increase in the reaction velocity, the rate ought to increase uniformly with the temperature. Such, however, is not the case, and this shows that chemical reactions are accelerated by a rise of temperature from more than one cause. Unfortunately, up to the present time, no satisfactory explanation of the enormous effect of temperature on the velocity of reactions has been given.

It is an interesting fact that the acceleration of many vital processes produced by rise of temperature is very similar to that observed for ordinary chemical reactions. For example, the rate of growth of yeast cells, the rate of germination of certain varieties of seeds, and the rate of development of the eggs of certain fishes is doubled or trebled for a rise of temperature of 10° C., which suggests that these processes are mainly chemical.

8. Influence of the Medium on the Velocity of Reaction.—Another factor which has a marked influence on the velocity of a chemical reaction is the nature of the solvent. This fact has been very clearly proved by the experimental work of Menschutkin.[2] A few of his results for the action of the trimethylamine on ethyl iodide, in different solvents, are given in the accompanying table, in which k denotes the velocity constant.

VARIATION OF VELOCITY CONSTANT WITH REACTION MEDIUM

Medium	k	Medium	k
Hexane................	0.00018	Methyl alcohol...........	0.0516
Ethyl ether.............	0 00076	Acetone.................	0.0608
Benzene................	0.00584	Benzyl alcohol...........	0.1330
Ethyl alcohol.	0.0366		

Note that the velocity of reaction in benzyl alcohol is about 740 times that in hexane. So far, no satisfactory explanation of the great differences in the velocities in different solvents has been furnished.

9. Rôle played by Catalysts in Changing the Velocity of Chemical Reactions.—The reaction of a catalyst may be pictured in very much the same manner as the action of heat, in that the driving force of an existing chemical reaction is accelerated, or the resistance of the reaction is reduced. Sometimes both effects are produced through the introduction of a catalyst, in which case the reaction is very much accelerated. Most catalysts, however, are considered to effect their acceleration of chemical reactions by diminishing the resistance. A model may make this clearer. Suppose that a sheet of metal is placed at such an angle that a weight placed at its upper end very slowly slides to the foot. The momentum of the moving body may even be insufficient to carry it all the way to the bottom. If a trace of oil be placed either on the plate of metal or on the bottom of the weight, the weight will slide *rapidly* to the foot of the plane. The oil remains unchanged. No energy has passed from the oil to either the weight or the plane, yet the rate of falling and the point of equilibrium have been altered. The oil may be taken as representing a catalyst. Burns says: "A catalyst is like a tip to a waiter, in that it accelerates a reaction that otherwise would proceed with infinite slowness."

[2] Zeit. phys. Chem., **6**, 41, 1890.

10. Characteristics of Catalytic Actions.—Catalysts are of very many kinds, and the mechanism of their action is so varied and so complicated that only a few general principles can be presented.

(a) *A catalyst is usually present in relatively small concentrations.*—One of the most striking and important features exhibited by catalytic phenomena is the very small quantity of the catalyst that is necessary to change the velocity of a chemical reaction. An illustration of this fact is furnished by the influence of molybdic acid on the rate of the reaction between hydrogen peroxide and hydriodic acid. It has been found that the velocity of the interaction between these substances in 0.01 N solution is more than doubled by the addition of molybdic acid in a concentration of 1 gram-molecule in 1,000,000 liters.

(b) *There are both positive and negative catalysts.*—If we accept Ostwald's definition of a catalyst—"A catalytic agent is that material which affects the velocity of a chemical reaction without appearing in the final products of the reaction"—we must conceive that a catalyst may either *accelerate* or *retard* a reaction. A catalyst that exerts an accelerating influence is known as a *positive catalyst*; whereas a catalyst that exerts a retarding influence is called a *negative catalyst*. The great majority of catalysts, however, belong in the positive class; there are only a few that are definitely known to retard certain chemical reactions.

Water vapor retards the rate at which phosphorus undergoes oxidation.

Chlorine retards the rate at which ozone is produced.

Oxygen retards the rate at which hydrochloric acid is formed.

(c) *Catalysts are unchangeable.*—The most striking characteristic of all catalysts is the fact that they have the same chemical composition at the end as at the beginning of the reaction. The amount of the catalyst is likewise unchanged.

In some cases the catalyst is found to emerge from the reaction in an altered physical state. For example, the crystalline variety of manganese dioxide, which acts as a catalyst in the decomposition of potassium chlorate, is found at the end of the reaction in the form of a fine powder.

(d) *It is doubtful whether catalysts are capable of starting a reaction.*—It is still a moot question among scientists whether or not a catalyst is capable of initiating a chemical reaction. Take, for example, the reaction between hydrogen and oxygen. If these gases are kept together at room temperature for an almost indefinite time, no

apparent combination takes place, unless some catalyst, such as finely divided platinum, is added. Since combination can also be made to take place rather rapidly by merely raising the temperature, Ostwald contends that the reaction does actually take place at room temperature, but that the reaction velocity is too small to admit of measurement. Other scientists claim that no reaction occurs at room temperature, and that the catalyst does actually initiate the combination of the gases.

(c) *Catalysts do not influence the final state of equilibrium.*—If a catalyst merely alters the rate of speed of a chemical reaction, and is not combined with any of the products when the reaction is complete, it does not modify the energy content of the system. If this is so, the final state of equilibrium of the reactive system must be independent of the catalyst. This statement has been proved experimentally. Take, for example, the case of a reversible reaction. In this instance the position of equilibrium is defined by the value of the equilibrium constant, and, as has already been shown, this constant is equal to the ratio $\dfrac{k_1}{k_2}$, where k_1 is the velocity coefficient of the direct reaction, and k_2 is that of the inverse reaction. Now, if the equilibrium is not disturbed by the introduction of a catalyst, it follows that the velocities of the two reactions must be accelerated in the same proportion. This fact has been proved in connection with the catalytic action of acids on the velocities of the hydrolysis of an ester and esterification.

(f) *Some foreign substances tend to retard, others to augment, catalytic action.*—The activity of a catalyst may be appreciably altered by the presence of even traces of certain foreign substances. Substances that tend to retard the action of a catalyst are known as *poisons*, whereas those that tend to augment the activity are known as *promoters*. The classical instance of a catalytic poison is found in the manufacture of sulphuric acid by the contact process, where it was found that the platinum asbestos employed for the oxidation of sulphur dioxide deteriorated so quickly as to render the process unworkable. Fortunately, it soon developed that the premature "ageing" of the catalyst was due to the presence of traces of *arsenic* in the sulphur dioxide. The arsenic poisoned the catalyst.

It is interesting to note that arsenic tends to paralyze the action of most catalysts. The same statement applies to prussic acid, iodine, mercury bichloride, etc.

Biological catalysts, or enzymes, are also subject to an analogous inhibiting influence, often, in fact, by the same substances that "poison" strictly inorganic catalysts.

In contradistinction to the above-mentioned inhibiting substances, it has been found that there are other substances which, when present in minute traces, greatly augment the activity of a catalyst. For example, finely divided oxides greatly increase the activity of metallic catalysts.

(g) *The activating power of many catalysts is often specific.*—According to Ostwald, every chemical reaction is more or less susceptible to catalytic influence, and there is a possibility that every substance acts in some degree as a catalyst. Yet each chemical reaction is especially influenced by some particular catalyst, or group of catalysts. In fact, some catalysts are very specific in their action, often showing but little effect on one of two closely allied reactions though acting on the other. Thus, hydriodic acid is slowly oxidized by hydrogen peroxide with separation of iodine. This reaction is catalyzed by tungstic acid, but the closely related reactions of the oxidation of hydriodic acid by persulphates or bromic acid are not affected by tungstic acid. In general, very pronounced catalytic activity is usually specific.

11. Autocatalysis.—Some reactions are catalyzed by the presence of one or more of their own reaction products and are, therefore, said to be *autocatalyzed*. Thus, when metallic copper is dissolved in nitric acid, the reaction proceeds slowly at first, then a little faster, and finally the speed of the reaction is greatly augmented. The acceleration is due to the catalytic action of the oxide evolved. Again, when a neutral ester is hydrolyzed by water, free acid is produced, and the acid acts as a catalyst.

Catalytic activities are also often "poisoned" by the action of autocatalytic substances.

12. Water as a Catalyst.—The presence of water seems to be absolutely essential for the occurrence of many reactions. Thus, it has been shown that no reaction occurs when hydrochloric acid and ammonia are brought together in a perfectly dry condition, but that when a minute trace of water is introduced, the reaction proceeds with great speed. The mechanism of this reaction is not thoroughly understood, but it is probable that the few drops of water effect chemical combination by starting ionization.

13. Hydrogen and Hydroxyl Ions as Catalysts.—Hydrogen and hydroxyl ions act as catalysts for many reactions which occur in aqueous solution. The reaction velocity of substances catalyzed by either of these ions is, in dilute solution, proportional to the concentration of the ions in question, provided the thermodynamic environment remains constant. The ion probably acts as a carrier, producing an unstable perhydrate as an intermediate product.

14. Catalysis in Heterogeneous Systems.—The catalytic reactions with which biological chemists are mostly concerned are those in which the catalyst occupies a phase distinct from that of the substrate. The reactions that are catalyzed by colloidal catalysts belong to this class. All enzymes are colloids; hence, all biologically catalyzed systems are heterogeneous in nature. Often, even the substrate is colloidal.

Colloids, as we have seen, are characterized by the development of surface, and adsorption is a consequence of surface phenomena; i.e., the intensity of adsorption is mainly dependent on the area of the adsorbing surface.

When one substance is adsorbed on the surface of another, it does not follow of necessity that any chemical reaction will occur; when it does take place, however, the rate at which it proceeds is obviously controlled by the amount adsorbed at any given moment. A relation between the concentration and the velocity of reaction is therefore to be expected.

The increased concentration of reacting substances at the surface of an adsorbing catalyst will in itself mean an increased reaction velocity, but this factor alone is not enough to account for the enormous increase which takes place in some instances.

When pronounced catalytic action takes place at a surface of contact, it is known as *contact catalysis*. This type of catalysis is shown by a number of metals and other solids, when present in solution in the colloidal state; and many important industries are founded on this fact, for example, in the manufacture of sulphuric acid, the synthesis of ammonia, the Deacon chlorine process, the Ostwald nitric acid process, the Sabatier hydrogenation process, and Welsbach's incandescent process (gas mantle).

The inorganic colloidal catalysts are of especial interest to biological chemists, because they offer the most striking analogy to the enzymes. They are even "poisoned" by the same substances that poison

enzymes. This point has been especially studied by Bredig and his pupils,[3] who made an elaborate study of the action of "poisons" on the catalytic activity of colloidal platinum.

Bredig, however, was careful to point out that the analogy is only an analogy. He does not think there is any identity between the action of the two classes of substances, which are themselves so different. To quote his own words: "All these facts point to an unmistakable analogy between the contact actions in the inorganic world and the actions of ferments in the organic word. As, in the case of my colloidal catalyzers, we are dealing with reactions in which enormously developed surfaces are involved, so it is probable that the same condition obtains in the actions of ferments, enzymes, blood corpuscles, and oxidizing and catalyzing substances. We see, therefore, that the organism develops its enormous surfaces in the tissues and colloidal ferments not only because it requires osmotic processes, but on account of the very great catalytic activity of such surfaces. If, as Boltzmann says, the war for existence which living matter must wage is a war about free energy, certainly of all the forms of free energy the *free energy of surface* is the most important for the organism."

15. Mechanism of Catalysis.—No satisfactory general explanation of catalytic action has been given up to date. This is not to be wondered at, since there is little likelihood that all catalytic accelerations are due to the same cause. It is much more probable that the mechanism of acceleration varies with the nature of the catalyst and with that of the reacting substances. In the case of homogeneous or one-phase catalytic systems, the most generally accepted explanation is that the catalyst reacts chemically with one of the reacting substances to form an intermediate compound, which in turn reacts in such a way as to regenerate the catalyst and produce the final products of the reaction. Thus, if a reaction of the type

$$A + B = AB$$
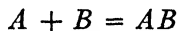
takes place very slowly under the ordinary conditions, it may be possible to accelerate the velocity of the reaction by the addition of a catalyst, C. In this case, according to the intermediate-compound theory, the reaction would proceed in the following two stages:

(1) $A + C = AC$,

(2) $AC + B = AB + C$.
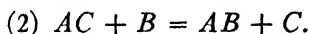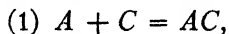
[3] Zeit. phys. Chem., **37**, 1, 1901; **37**, 323, 1901; **38**, 122, 1901.

Note that the catalyst is regenerated in the second stage of the reaction.

Catalysts that react in this manner are known as carriers. Thus, nitric oxide, NO, acts as a carrier in the well-known chamber process for the manufacture of sulphuric acid. The combination may be represented as taking place as follows:

$$(1) \quad 2NO + O_2 = 2NO_2,$$

$$(2) \quad NO_2 + SO_2 + H_2O = H_2SO_4 + NO.$$
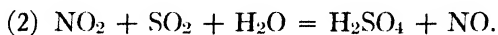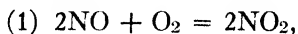
This explanation of the action of the oxides of nitrogen was suggested more than a century ago by Clement and Desarmes, and it is still considered the most plausible one although it does not tell the whole story.

In the case of heterogeneous catalysis, it seems probable that adsorption on the surface of the catalyst is an important part of the procedure. Nernst[4] has pointed out that there are three stages in heterogeneous reactions: diffusion, adsorption, and chemical reaction. The reacting bodies, e.g., enzyme and substrate, must first of all come into contact; hence, their rate of diffusion is the *first* factor. The next is that of adsorption, which usually takes place very rapidly. The final process, the chemical change, is determined by the amount of substrate adsorbed at a given moment, in accordance with the law of mass action. For further information on the mechanism of catalytic action, the student is referred to an elaborate article by Bancroft,[5] in which are presented the various theories that have been advanced in explanation of the phenomenon.

16. Enzymes as Catalysts.—From what we have just learned in regard to catalysis as a concept, it is evident that practically every chemical reaction is concerned to a greater or less extent in the phenomenon. The reactions occurring within living organisms are no exception to the rule; in fact, it would be difficult to find a biochemical reaction that is not influenced by some catalyst, and usually the catalyst is one of biological origin. Thus, before the various food-stuffs can be utilized for the building up of new tissue, they must be hydrolyzed into simpler compounds, and these hydrolytic processes are hastened by the action of enzymes. The hydrolytic, oxidative, and

[4] Zeit. phys. Chem., **47**, 52, 1904.
[5] Trans. Am. Electrochem. Soc., **87**, 233, 1920.

reductive reactions which take place within the living cells of plants and animals are likewise catalyzed by enzymes.

Provisionally, we may define an enzyme as an organic catalyst that is produced by an animal or vegetable cell.

Some enzymes act within the cells, while others are secreted by the cells and act on a substrate outside the cell. The former are known as *endoenzymes*; the latter, as *ecloenzymes*.

The action of enzymes closely resembles that of the inorganic catalysts, but there are certain differences, presumably depending on their constitution and on their peculiar surroundings.

In the case of inorganic catalysts, we have seen (p. 323) that very small quantities can greatly affect the speed of a chemical reaction, and the same holds true for enzymes. In fact, it is usually found that in reactions which can be enhanced either by an enzyme or by an inorganic catalyst, the enzyme action is the more powerful of the two. For example, in the hydrolysis of lactose into glucose and galactose, the enzyme *lactase* is about 5000 times as effective as an equal weight of hydrochloric acid.

Like inorganic catalysts, the enzymes do not modify the nature of the final products of a reaction; they merely change the rate of speed of the reaction, and they do not appear in the final products. Some enzymes, notably diastase, can be recovered in an unaltered state and amount after they have completed their action. Enzymes, therefore, are true catalysts.

17. The Chemical Nature of Enzymes.—Although much work has been done on the chemistry of enzymes, we are still ignorant in regard to the chemical structure of enzymes, owing to their tendency to coagulate and precipitate when they are subjected to the action of heat or strong chemicals. Their existence is recognized in a solution by their action. They are generally regarded as non-diffusible colloidal particles. Owing to the fact that some of the most highly purified enzymes that have been isolated, notably pepsin, malt amylase, pancreatic amylase, and lipase, give the reactions characteristic of proteins, many authors have advanced the idea that enzymes are peculiarly endowed proteins. The question is still an open one, however. Quite recently, Euler and Josephson[6] obtained highly purified and very active preparations of invertase (an enzyme which hydrolyzes sucrose to glucose and laevulose). Their products gave undisputable protein

[6] Ber., 56 B, 1097, 1923; *ibid.*, 57 B, 299, 1924; Zeit. phys. Chem., **133**, 279, 1924.

reactions. On the other hand, Willstätter and Kuhn[7] have produced an active preparation of invertase which is free from protein, carbohydrate, and phosphorus. It is evident, therefore, that the chemical nature of the *active* principle of invertase is still an unknown quantity.

Attempts to isolate an enzyme in crystalline form have generally been met with failure. Quite recently, however, the feat has been accomplished. Sumner[8] reports that by refrigeration of an acetone extract of jack-bean meal he obtained a product of sharply defined octahedral crystals, which he believes to be identical with the well-known enzyme urease (an enzyme which decomposes urea to ammonium carbonate). The product is protein in nature and is presumably a globulin.

The action of an enzyme, which alone furnishes the proof of its existence, can be investigated only by discovering the nature of the compounds that are formed, and, when these are known, by studying the rate of the change and the various conditions which affect it.

18. The Specific Nature of Enzyme Action.—Most enzymes are highly specific in their action, whereas inorganic catalysts are less so. Enzymes usually act only on one compound, or type of compounds; e.g., invertase acts only on cane sugar, maltase acts only on the alpha-glucosides, emulsin (from almonds) acts only on the beta-glucosides, the enzyme that acts on lactose will not act on either sucrose or maltose, and so on. That specificity of action is also found among the inorganic catalysts has already been pointed out (p. 325), but it is not so well marked as it is with the enzymes. One cannot therefore regard the limitation of action of enzymes as a peculiar property; nevertheless, when all deductions have been made, the highly specific nature of enzymes is an outstanding fact which has to be accounted for. The enzymes that hydrolyze the fats do not hydrolyze the proteins, and the same is true of the enzymes that act on carbohydrates.

It must not be inferred that each individual substance has its own specific enzyme for every kind of chemical reaction. The lipases act on many simple organic esters, and maltase acts not only on maltose but on other alpha-glucosides, etc. It is clear, therefore, that enzymes exhibit a group specificity rather than a substance specificity.

The phenomenon of specificity of action is explained by supposing that a union occurs between the enzyme and the substrate, and in

[7] Zeit. phys. Chem., **125**, 28, 1923.
[8] J. Biol. Chem., **69**, 435, 1926.

order that this union may take place it is assumed that the enzyme possesses a configuration which corresponds accurately with that of the substrate. Emil Fischer has compared the action of enzymes to that of a lock and key where only one key will open a particular kind of lock, no other key will fit it, nor will its key fit other locks.

The phenomenon of specificity of action is one of great importance in physiology, since it explains how a given chemical reaction taking place within a living cell may be augmented or retarded without affecting the rate of any other chemical change in the same cell or vicinity.

19. Reversibility of Enzyme Action.—Enzymes resemble inorganic catalysts also in that, where the reaction involved is a reversible one, they augment both the forward and the reverse changes. This is just what we should expect, for an enzyme, like any other catalyst, merely alters the rate of a chemical reaction (p. 324), and hence must shorten the time necessary for the establishment of equilibrium, whether it be a direct or reverse reaction. "Since a chemical equilibrium is dynamic and is produced by the velocity of the reverse reaction becoming equal to the velocity of the direct reaction, it follows that if the position of equilibrium is not altered by an enzyme, the latter must accelerate the reverse and the direct reactions equally." (Findlay.)

This fact has been proved in a variety of cases. An instance is furnished by the action of lipase on the esters of the lower fatty acids. Kastle and Loevenhart[9] have shown that if this enzyme is allowed to act on ethyl butyrate in the presence of water, partial hydrolysis into butyric acid and ethyl alcohol takes place; while if it is allowed to act on an aqueous mixture of butyric acid and ethyl alcohol, a certain amount of the ester will be produced. Ethyl butyrate is an excellent compound to use in illustrating this phenomenon; on account of its odor, the ester is easily recognized.

20. Synthetic Reactions of Enzymes.—In general, the course of a reaction induced by an enzyme will conform to the law of mass action. It is evident, then, that the concentration of the substrate is an important factor in determining whether *synthesis*, the building up of more complex compounds, or *cleavage*, the breaking down of compounds, will tend to take place. Thus, in a 10 per cent solution of maltose containing the enzyme maltase, hydrolysis proceeds until about 98 per cent of the maltose is converted into glucose, whereas in

[9] Am. Chem. J., **24**, 491, 1900.

a 40 per cent solution equilibrium is established when about 85 per cent of the maltose has undergone hydrolysis. Croft Hill[10] showed that if maltase was added to a 40 per cent solution of glucose, a change took place in the reverse direction; in other words, a synthesis of some maltose took place, and this reaction continued until 85 per cent of the glucose was left.

The reverse action of an enzyme, however, on the products of a reaction may not be absolutely the reverse of its affect on the direct reaction, for it has been found that the sugar formed through the action of maltase on glucose is an isomeric form of maltose, called, *iso-maltose*. Since the discovery of this synthetic action of an enzyme by Croft Hill, "so many other cases have been discovered that the impression is distinctly given that it is merely a question of finding the proper conditions to be able to obtain synthesis from all enzymes " (Bayliss).

Most enzyme reactions take place in the presence of excess of water, and hence it follows that hydrolysis predominates over synthesis. In synthetic reactions the amount of water must be limited, but in all cases water must be present. For example, pure glycerol, fatty acids, and lipase will not react to form fat until a little water is added.

The foregoing facts regarding disintegration and synthesis are of great biological interest, since they help us to understand more clearly the reactions which occur in the living cell; reaction which, *in vitro*, can be carried out only by means of strong reagents and at high temperatures, yet are performed with perfect ease in the living organism.

21. Classification and Nomenclature of Enzymes.—Owing to the fact that practically nothing is known in regard to the chemical composition of enzymes, they can only be studied by considering their effect on various substrates. In the first place, with regard to nomenclature, the earlier workers, supposing them to be proteins, gave them names ending with the suffix *in*, similar to those of the proteins   Many of these names are still used, e.g., pepsin, trypsin, ptyalin, pancreatin, etc.; but since there is considerable doubt that all enzymes are protein in nature, it has become customary to name the enzymes according to the substances on which they act. This is done either by replacing the last part of the name of the substance acted upon by the suffix *ase*,

---

[10] J. Chem. Soc., **73**, 634, 1898.

e.g., sucrase, or by merely adding *ase* to the name of the substance acted upon, e.g., glycogenase.

The substance upon which an enzyme acts, or, to be more exact, the substance whose hydrolysis, oxidation, reduction, or other chemical change is catalyzed by an enzyme, is called the *substrate.*

Most of the enzymes accelerate hydrolytic reactions, and are spoken of as "hydrolases." Those that accelerate oxodative processes are called "oxidases" and "peroxidases"; while those that affect the rate of speed of reduction are called "reductases"; those that catalyze the splitting off of amino-acid groups are called "deaminases"; and those that assist in the splitting off of CO_2 from COOH groups are called "carboxylases," etc.

22. Combination of Enzymes with Substrates.—"There is abundant evidence that a combination of some kind is formed between the enzyme and the substrate preparatory to the action of the former. There is also similar combination between enzyme and products" (Bayliss).

All the facts bearing on the specificity of enzyme action confirm this conclusion, for these facts can only be explained on the hypothesis that enzymes form addition compounds with the various substrates.

The nature of the compound formed between the enzyme and the substrate is at first, at any rate, an adsorption compound, since it is found that the laws of adsorption are obeyed, and we have very definite evidence of the formation of such compounds. For example, the enzyme amylase, which hydrolyzes starch, can pass through a porous clay filter by itself, but if it is added to a 2 per cent starch solution it cannot pass through until after the hydrolysis takes place. Adsorption of enzymes with substrates on which they have no catalytic effect may also take place; e.g., amylase is adsorbed by calcium caseinogen although it does not act on it. Again, enzymes are much more sensitive to heat when in a relatively pure state than when in combination with the substrate; e.g., sucrase will stand a temperature 25° C. higher in the presence of sucrose than when pure. These facts can only be explained by assuming that adsorption takes place between the enzyme and the substrate.

23. The Effect of Temperature on Enzyme Action.—We have already learned that a rise in temperature increases the velocity of a chemical reaction very considerably. In general, in *homogeneous* systems the velocity of reaction is doubled or trebled for each increase

in temperature of 10° C. This rule applies to chemical processes catalyzed by enzymes only over a very limited range of temperature. In other words, an *optimum temperature* for each enzyme is soon reached, beyond which, as the temperature rises, the reaction is first slowed down and then ceases altogether.

This peculiarity of biological catalysts as compared with inorganic catalysts does not disprove the analogy between the two types of catalysts. As a rule, enzymes, being colloidal in nature, are coagulated by heat, or precipitated, and hence rendered inactive. Often, not only the enzyme, but the substrate as well is colloidal in nature and therefore sensitive to changes in temperature. As stated above, the optimum temperature is that at which the velocity of an enzyme action is at a maximum. The optimum temperature of most enzymes lies between 35° and 55° C. At temperatures slightly above the optimum alterations occur in the physical state of the enzyme, and in the substrate if it is likewise colloidal. These alterations in physical state, such as viscosity and conductivity, indicate an increase in the size of the colloidal particles, and consequently a decrease in their specific surface. This implies a reduction in the effective adsorbing surface. At the optimum temperature the catalytic effect of the enzyme and the increased chemical action due to temperature are at a maximum. (In homogeneous systems, still higher temperatures would be more effective.) Beyond this temperature, the loss of specific surface due to coagulation becomes an important factor, and finally a point is reached at which the adsorbing power is totally lost, chemical action ceases, and the enzyme is said to be dead. The "killing" temperature of most enzymes is between 60° and 80° C.; but a few are known that can withstand a temperature of 100° C., if they are well protected by the substrate or products of reaction, and are in a favorable reaction medium.

Extreme cold has no effect, except to slow down the process. Enzymes are not destroyed by exposes to the temperature of liquid air for several days.

24. Enzyme Activity and Hydrogen-ion Concentration.—The activity of enzymes is greatly dependent on the hydrogen-ion concentration of the medium; in fact, for each enzyme there is a definite pH at which the maximum activity of the enzyme is exhibited. In many cases a very slight deviation from the optimum pH has a marked influence on the activity of the enzyme; and in all cases it is

only when the medium has the optimum pH that constant values are obtained for the velocity coefficient of the reaction. This is clearly shown by the work of Sörensen,[11] Northrop,[12] and others. In Fig. 38 is represented the relation between the activity of an enzyme and the concentration of hydrogen ions, the ordinates representing the percentage of substrate hydrolyzed, and the abscissæ the pH values. In this case the substrate is sucrose, and the enzyme sucrase. The two
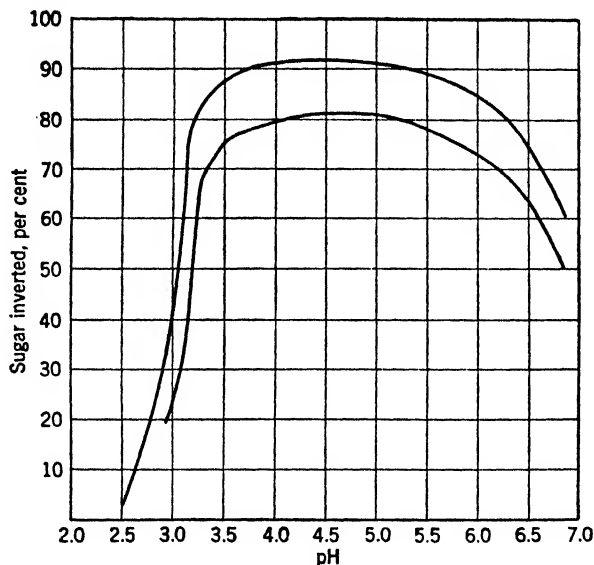
FIG. 38.—Influence of pH on Enzyme Acitivity. (After Sörensen.)

curves refer to experiments with two different solutions of sucrase. It will be noted that the optimum enzyme activity is obtained when the pH equals 4.4–4.6.

After performing a long series of experiments, Northrop (*loc. cit.*) came to the conclusion that the determining factor in the hydrolysis of proteins by pepsin is the amount of ionized protein in solution. At the iso-electric point of a given protein, the action of pepsin is practically *nil*, showing that it is the protein ion which is acted upon by the enzyme. Furthermore, the optimum pH for the activity of pepsin coincides with the pH at which the protein solution is most highly ionized.

[11] Biochem. Zeit., **21**, 131, 1909.
[12] J. Gen. Physiol., **8**, 211, 1920.

In the following table are given the values of the optimum pH for a number of enzymes:

Enzyme	Optimum pH
Sucrase, yeast	4.4–4.6
Trypsin (hydrolysis of peptone)	8.3
Pancreatic lipase	8.0
Rennin	5.7
Erepsin	7.7
Amylase, pancreatic	7.0
Amylase, saliva	6.0
Catalase (blood)	7.0
Maltase	6.6
Pepsin (hydrolysis of egg albumin)	1.6

It is evident from the foregoing data that one who is studying enzyme reactions should work with solutions of regulated pH.

25. The Ionic Nature of Enzymes.—The very marked effect of the hydrogen-ion concentration of a solution on the activity of an enzyme suggests that many, if not all, active enzymes are ionic in nature. Northrop[13] has recently made an important contribution to this phase of the subject. His work is based on the Donnan equilibrium, which as we have seen (p. 295), is a reliable index of the ionic nature of any diffusible substance, since the ratio of concentration of any ion on the two sides of a membrane must be equal to the ratio of the concentrations of any other ion of the same sign or valence, whereas a non-ionic substance would be equally distributed on both sides. Northrop applied this principle in a study of the distribution of trypsin inside and outside of gelatin particles, at different hydrogen-ion concentrations, and compared it with the distribution of chlorine and hydrogen ions under similar conditions. He found that trypsin behaves like a monovalent positive ion from pH 2 to 10.2. At this point it behaves as though it were un-ionized, and on the alkaline side of pH 10.2 it becomes a monovalent negative ion. From this we infer that trypsin is probably iso-electric at pH 10.2.

26. Quantitative Relationships in Enzyme Reactions.—We have seen that when a substance is hydrolyzed by acids the relationship between the amount of unaltered material in the system and the velocity of change is rather a simple one; for example, in the hydrolysis of sugar, the amount of unaltered sugar and the velocity of decom-

position stand in a constant ratio to one another throughout the reaction. In the case of most enzymatic reactions, however, it has been found that when the velocity is studied quantitatively the ordinary mass-action equation is followed only when the concentration of the substrate is not too great. In other words, when the quantity of substrate material becomes smaller than that which permits the maximum combining power of the enzyme to take effect, the reaction immediately slows up. Thus, when invertase is allowed to act on fairly concentrated solutions of sucrose, all of the invertase is in combination with the sugar, and the portion of the sugar that is thus combined is the only portion undergoing perceptible hydrolysis. For a given concentration of invertase, therefore, the rate of hydrolysis of sugar is constant; and for varying amounts of invertase, the velocity of reaction is in direct proportion to the amount of enzyme present. Mathematically expressed:

$$\text{Velocity} = kE,$$

where k is a constant and E is the concentration of the enzyme. This condition will continue as long as there is an excess of substrate above that required for *adsorption-saturation* of the enzyme, for under these conditions the amount of sucrose adsorbed on the enzyme will be constant and the velocity of reaction must also remain constant. That is, the same amount of substrate will be acted upon in unit time. As the reaction proceeds, however, a point must be reached at which the amount of substrate present will no longer be sufficient to bind all of the enzyme, and the velocity of reaction will then become proportional to the amount adsorbed. After that the velocity will gradually die off until equilibrium is established.

In many cases, as the concentration of the substrate diminishes below the adsorption-saturation point, other factors tend to render the velocity coefficient inconstant; i.e., it will either increase or decrease as the reaction proceeds. This inconstancy in the velocity is usually due to the accumulation of the products of the reaction in the system. Sometimes the products of the reaction tend to augment the velocity coefficient by supplying an "activator," while in other cases the velocity is decreased by the production of a "poison," or, more often, by formation of an enzyme-product adsorption compound which is inactive. In those cases in which acids are produced as one of the products of the reaction, the velocity is at first accelerated, but

finally, as the acid increases, it brings about the destruction of the enzyme.

It is evident from the foregoing data that, if one wishes to determine the velocity of an enzyme, all factors, such as concentration of substrate, hydrogen-ion concentration, temperature, etc., must be kept constant, the amount of enzyme alone being altered.

27. The Effect of Electrolytes on Enzyme Action.—It is evident than any foreign substance may affect the kinetics of enzyme reactions. The presence of electrolytes, for example, may have considerable effect, enzymes being very sensitive to the action of electrolytes. Acids and alkalies are, of course, especially powerful in their effect.

A distinction must be drawn at the outset between the more general action of electrolytes and the *specific* action of certain electrolytes on particular enzymes. These will be discussed later under the heading of "co-enzymes."

The effect of acids and alkalies, which is attributable to changes in pH, has already been discussed (p. 334).

Neutral salts also tend either to increase or to decrease the velocity of enzymotic reactions, but their action hardly admits of classification. Many cases are known in which low concentrations of certain electrolytes will change the velocity of particular enzyme actions, by enhancing the ionization of the substrate, the enzyme, or both; but in high concentrations *all electrolytes* inhibit the action of enzymes.

28. Co-enzymes.—Many enzymes have been found to be practically inactive unless they are accompanied by some other substance, which may be an electrolyte or a non-electrolyte. Phosphates are required in the case of salivary amylase, and the bile salts in the action of pancreatic lipase. Manganese is an essential constituent of most oxidases. Such substances are known as co-enzymes.

Another well-known example of a co-enzyme is found in alcoholic fermentation. Yeast cells contain an enzyme, or rather an enzyme system, which brings about the conversion of sugar into alcohol and carbon dioxide. It has been found that if an aqueous solution of yeast is filtered through a Martin's gelatin filter, which keeps back colloids, neither portion is active by itself, but if the two portions are mixed, the mixture is active. Since inorganic phosphates increase the activity of yeast juice, it was thought that the phosphate ion might be the co-enzyme of zymase, but experiments showed that the phosphate ion alone is not sufficient to restore the activity; some other

substance in the filtrate, whose exact nature has not been determined as yet, is also necessary. In short, both substances are required.

We have already seen that pepsin is inactive except in the presence of free acid, and trypsin is inactive except in the presence of alkali. These two cases are probably due to ionization.

29. Paralyzers and Anti-enzymes.—It has already been pointed out (p. 324) that enzymes, as well as inorganic catalysts, are very sensitive to a large variety of substances, both organic and inorganic, traces of these substances rendering them inactive. In the case of enzymes, these "inhibitors" are of two classes: "paralyzers" and "anti-enzymes." Among the paralyzers are the salts of the heavy metals, which have a tendency to alter the physical state of colloids in general, and therefore, of colloidal enzymes. Formaldehyde, hydrocyanic acid, fluorides, and a number of other substances commonly used as antiseptics, inhibit not only enzyme action, but also that of the living cells which contain the enzyme. Certain other antiseptics such as toluene, xylene, etc., inhibit the growth of the cell, but do not affect the action of the enzymes which may be present. It is evident, therefore, that by the careful selection of antiseptics one can often distinguish between the chemical changes that are involved in the life process of a cell and those that are specifically enzymotic.

The "anti-enzymes" are inhibitors of biological origin. They probably combine with the enzyme and thus prevent it from acting on the substrate. These anti-bodies are very widely distributed, as can be seen from the fact that anti-enzymes have been discovered for rennin, pepsin, trypsin, lipase, urease, amylase, lactase, tyrosinase, emulsin, papain, and thrombin. It is generally assumed that the reason why the stomach does not digest itself is that during life an anti-enzyme (anti-pepsin) is present in the mucous membrane of this organ. A similar anti-body (anti-trypsin) is found in the intestinal mucosa and in intestinal worms.

30. Zymogens.—" Since all enzymes are produced by the agency of living protoplasm, it is evident that at some stage in their formation they must be devoid of the catalytic properties of the fully formed enzyme. This stage is called a "zymogen" when it can be obtained free from the cells in which it is formed and can be converted by purely chemical means into the active enzyme" (Bayliss). As far as we know, enzymes are not reconvertible into the zymogenic form.

This phenomenon has been studied especially in connection with

the enzymes pepsin and trypsin, both of which are obtainable in the zymogenic state. "Pepsinogen" in converted into active pepsin by acids, especially hydrochloric acid, which is also manufactured in the stomach mucosa. Trypsinogen, which is secreted in the pancreatic gland and flows into the duodenum in this state, may be rendered active by contact either with calcium salts or with a co-enzyme, enterokinase, which is secreted in the intestinal tract.

31. Oxidation-reduction Reactions.—The energy of the various life processes is derived mainly, if not entirely, from oxidative reactions carried out by the living cells. The substances so oxidized are, in general, products derived from our foodstuffs with the assistance of the hydrolytic enzymes. These products are comparatively stable in the presence of atmospheric oxygen, but they are rapidly oxidized in the animal body or plant tissues or in the presence of certain tissue extracts.

Scientists have long been impressed by the speed and ease with which these oxidation processes take place at the relatively low temperatures of the animal body and the still lower temperatures of plants, because, when the same substances are oxidized outside of living organisms, one must apply a flame and the combustion is carried out at high temperatures.

Many theories have been proposed to account for the oxidation processes which take place in the living body. We shall consider only a few of these. Before doing so, however, it may be well to mention a few well-established facts concerning oxidation processes.

An oxidation reaction is most simply defined as one in which a given substance combines chemically with oxygen; but this is not the whole story, for, as a rule, it can be shown that oxidation reactions involve the simultaneous reduction of some other substance. An exchange of electrons is involved in the process.

Oxidation, instead of being a simple reaction, occurs in stages. Take, for example, the oxidation of a very simple substance, carbon monoxide, which is usually represented by the equation $CO + O = CO_2$. It can be shown by experiment that the following steps are involved in this process:

$$CO + HOH = HCOOH$$
$$HCOOH = CO_2 + H_2.$$
$$H_2 + O_2 = HOOH$$
$$HOOH = HOH + \tfrac{1}{2}O_2.$$
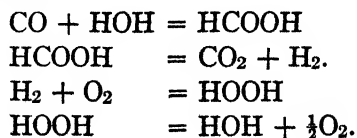
The reaction proceeds as indicated in the foregoing equations because of the presence of oxygen. The oxygen acts as a "hydrogen acceptor." But, before atmospheric oxygen can act in this manner it must be *activated*. At high temperatures the heat does the activating, whereas at body temperatures the activation is presumably effected through the agency of enzymes.

Many biological oxidations take place in aqueous solution in the absence of molecular oxygen. In these reactions, some easily reducible substance must also be present. The oxygen is obtained from the molecules of water; and some substance, termed the "hydrogen acceptor," undergoes reduction, while another, the "oxygen acceptor," is simultaneously oxidized. Under these conditions, oxidation becomes essentially a process of splitting up water into hydrogen and oxygen. If R is the substance oxidized and B is the hydrogen acceptor, we have, under anaërobic conditions,[14]

$$R + \begin{array}{c|c} OH & H \\ OH & H \end{array} + B \rightarrow R \underset{\underset{RO\ +\ HOH,}{\downarrow}}{\overset{OH}{\diagup}}_{OH} + B \overset{H}{\underset{H}{\diagup}}$$

Oxygen acceptor   Hydrogen acceptor
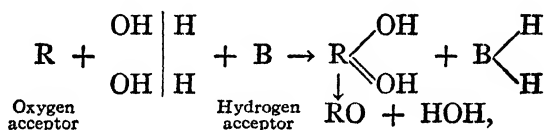
and, under aërobic conditions,

$$R + 2H \cdot OH + O_2 \rightarrow R(OH)_2 + HO \cdot OH \rightarrow RO_2 + 2H_2O.$$
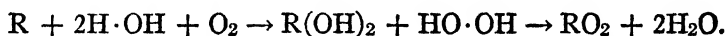
Wieland[15] has shown that many organic substances (hydroquinone, for example) can be oxidized in the complete absence of oxygen if certain catalysts are present. The oxidation, according to his theory, proceeds very largely by a primary loss of hydrogen under the influence of such catalysts as the oxidation enzymes, the hydrogen in turn being taken up by an acceptor, the oxidizing agent. As an illustration of this theory of oxidation, which is known as *oxidation by dehydrogenation*, we may take the oxidation of an aldehyde to an acid. In the presence of a suitable catalyst, such as palladium black or an oxidase, and of a hydrogen acceptor (oxidizing agent) such as methylene blue, the reaction may be represented as follows:

---

[14] Two kinds of oxidation are known to take place in living matter: that taking place at the expense of atmospheric oxygen; and that which is effected by easily reducible substances. The former is called *aërobic oxidation*; the latter, *anaërobic oxidation*.

[15] Ber., **46**, 3327, 1913; **54**, 2353, 1921.

$$RC\!\!\diagdown_{\substack{H\\O}} + \substack{H\\O\\H} \rightleftarrows RC\!\!\diagdown_{\substack{H\\OH\\OH}}$$

$$RC\!\!\diagdown_{\substack{\boxed{H}\\\boxed{OH}\\OH}} + \text{catalyst} + M \rightarrow RC\!\!\diagup^{O}_{OH} + \text{catalyst} + MH_2.$$
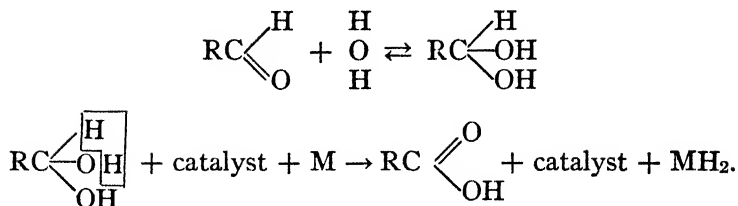
M represents a molecule of methylene blue, and MH_2 its reduction product, leuco-methylene blue. Note that here, as in the case of the oxidation of CO to CO_2, the first step in the oxidation is hydration; the second step is one of dehydrogenation.

Another theory of oxidation that has found considerable favor is that of *activation of oxygen*, the activation being induced or accelerated by a type of catalysts known as "peroxidases." This name is given to these catalysts because the oxidized substance forms a peroxide as an intermediate product. If we take the oxidation of an aldehyde to an acid we can represent the reactions as follows:

$$RC\!\!\diagdown_{\substack{H\\O}} + O_2 \rightarrow RC\!\!\diagdown_{\substack{O-OH\\O}}$$

$$RC\!\!\diagdown_{\substack{O-OH\\O}} + RC\!\!\diagdown_{\substack{H\\O}} \rightarrow 2RC\!\!\diagdown_{\substack{OH\\O}}$$
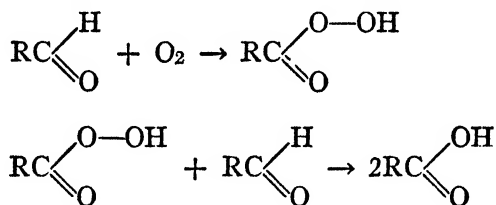
We have many such reactions taking place in the body.

32. Physico-chemical Explanation of Oxidation-reduction Reactions.—It is evident from the foregoing data that oxidation and reduction processes cannot be expressed by the simple reactions

$$X + O_2 = XO_2, \quad \text{or} \quad X + H_2 = XH_2,$$
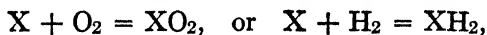
since before such reactions can take place the oxygen or the hydrogen must be "activated." Furthermore, it has been proved that oxidation involves the simultaneous reduction of some other substance and *vice versa*. In other words, the processes of oxidation and reduction are not merely opposite; they are also complementary, for, unless the oxidation takes place in the presence of free oxygen, e.g., in air, the oxygen required for the oxidation must be derived from some substance containing it, and the latter substance, in virtue of giving up oxygen, is

reduced at the same time as the oxidation takes place. Even the oxidations that take place in air are not direct, but take place in stages.

We shall now attempt to interpret oxidation-reduction reactions in the light of our modern conception of atomic structure. We have learned that the atoms of elementary substances are electrically neutral and exhibit neutral valence. When a substance like magnesium burns in oxygen, the essential change is the loss of 2 electrons by each magnesium atom and the gain of 2 electrons by each oxygen atom. In the product, magnesium exhibits a positive valence of 2, and oxygen a negative valence of 2. In other words, oxidation consists simply in the loss of one or more valence electrons by atoms or ions, and reduction in the reciprocal gain of electrons by atoms or ions. Take the following further examples: A hydrogen atom becomes oxidized by the removal of its electron and becomes a hydrogen ion with a valence of 1, which is its maximum valence. A cuprous ion, Cu^+, becomes oxidized by the loss of an electron, becoming the cupric ion, Cu^{++}, its valence changing from 1 to 2 by the oxidation. The ferrous ion, Fe^{++}, on losing an electron, becomes the ferric ion, Fe^{+++}, its valence changing from 2 to 3. The reaction between ferric chloride and potassium iodide, by which iodine is set free, is also an oxidation reaction, the iodide being oxidized by the ferric chloride. The reaction is as follows:

$$Fe^{+++} + 3Cl^{=} + K^+ + I^- \rightarrow Fe^{++} + 3Cl^{=} + K^+ + I.$$
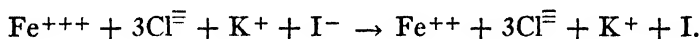
It will be observed that this oxidation reaction has really involved the transference of an electron from the ferric ion, which is the oxidizing reagent, to the iodine ion, which on acquiring the electron becomes oxidized to the atomic form. Another example of oxidation that does not involve oxygen is the oxidation of zinc by an acid. The essential change in this case is the loss of electrons by the zinc and the gain of electrons by the hydrogen ions. The reaction is as follows:

$$Zn_2 + \begin{matrix} H^+ + Cl^- \\ H^+ \quad Cl^- \end{matrix} \rightarrow Zn^{++} + \begin{matrix} Cl^- \\ Cl^- \end{matrix} + H_2.$$
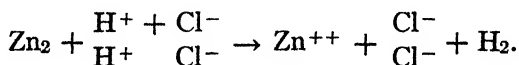
The hydrogen ion is the oxidizing agent since it removes electrons from the zinc atoms, the hydrogen becoming atomic and the zinc ionic. In other words, the hydrogen is reduced and the zinc is oxidized.

It is evident from the foregoing remarks that the activity of the so-called oxidizing agents depends on the tendency of some atom in the

molecule to gain electrons. Halogens are good examples of this type of bodies. In terms of the Langmuir theory of atomic structure (see Chapter II), the outer shell of the halogens requires only one electron to reach a stable configuration and they consequently show a great tendency to acquire these additional electrons. Molecules that contain such atoms readily receive electrons, because stable compounds can be formed when these elements are in a lower positive valence.

In general, then, the transfer of electrons from one substance to another, i.e., the oxidation of one substance by another, takes place at ordinary temperatures with perfect ease between free ions or atoms with exposed charges as compared with atoms of molecules that show little tendency to ionize.

Such compounds as the peroxides, RO—OR, persulphides, RS—SR, cyanogen, NC—CN, oxalic acid, etc., exhibit great activity because they contain an unstable atom in an intermediate condition of oxidation or reduction, and hence they possess a very strong tendency to go into a stable condition of greater oxidation or reduction by the loss or gain of electrons. In the peroxides the oxygen atom is in the unstable condition—($^+O^-$), and it readily loses an electron to become ($^-O^-$). The ($^+_+C^+_-$) in cyanogen and oxalic acid has a strong tendency to go over to the most stable condition of oxidation ($^+_+C^+_+$).

The rôle of catalysts is undoubtedly that of activators; that is, they assist in bringing about a condition of exposed electrons on some of the atoms of the compound on which they act. Colloidal platinum or palladium or oxidative enzymes augment the oxidation and reduction of comparatively inert substances presumably because of their high dissociating and ionizing powers, which lead to the dissociation or ionization and the accompanying exposure of charges on atoms in the molecules on which they act.

Cady and Taft[16] suggest the terms "electronation" for reduction and "de-electronation" for oxidation.

An "oxidase," then, augments the de-electronation of a substance; and a "reductase" assists in the electronation of a substance.

This chapter will close with a brief reference to two cases of oxidation that are of interest to students of the biological sciences.

H. Schode[17] has demonstrated that the oxidation of glucose by hydrogen peroxide is augmented enormously by the addition of a trace

[16] Science, **62**, 403, 1925.
[17] Die Physikalische Chemie in der Inneren Medizin, 226, 1923.

of iron sulphate. The iron salt evidently effects this transformation of energy as follows: In the first place, it activates the hydrogen peroxide, i.e., it effects the rapid transfer of electrons from the carbon atoms of the glucose to the oxygen of the hydrogen peroxide. This is a rather complicated reaction, involving the formation of ferrous peroxide which is quickly converted into ferric oxide, electrons passing from the ferro atoms to the oxygen of the peroxide.[18]

In the second place, the iron atoms activate the glucose by "de-electronation," just as the hydrogen peroxide was activated by "electronation," the iron acting as the *carrier* of electrons. Now, since the iron can be acted upon by both the activated glucose and the activated hydrogen peroxide, it is alternately reduced and oxidized; hence it is not destroyed by the reaction and therefore conforms to our definition of a catalyst.

A few years ago, Hopkins[19] made an epoch-making contribution to our knowledge of biological oxidations by isolating from plant and animal tissues an organic carrier of oxidation-reduction reactions. This auto-oxidizable substance is a derivative of crystein and was named *glutathione* by the discoverer. It is characterized by containing the sulph-hydryl group, SH, and is said to be subject to the reversible oxidation and reduction reaction indicated by the following equation:

$$2R\text{---}SH \rightleftarrows R\text{---}S\text{---}S\text{---}R + H_2,$$
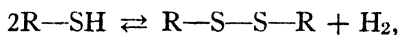
where R represents all but the sulph-hydryl group of the reduced glutathion.

There is considerable evidence to show that the real oxidation should be expressed thus:

$$2RS^- \rightarrow RS\text{---}SR + 2e,$$
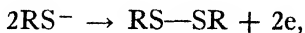
and that iron is the activator of the RS^- ion.[20] This sulphide ion readily gives up an electron, as is shown by the fact that acid retards the reaction, whereas a slightly alkaline reaction favors the action by increasing ionization. (See article by Dixon and Quastel.[21])

[18] Groad and Rideal, Proc. Royal Soc., **105**, 156, 1924.
[19] Biochem. J., **15**, 286, 1921.
[20] Warburg and Sakuma, Pflüger's Arch., **200**, 203, 1923.
[21] J. Chem. Soc., **123**, 2943, 1923

Whether or not the foregoing picture of the phenomenon of oxidation and reduction is correct in all particulars, it is unquestionable that the process does involve the transfer of electrons, and that this is the essence of it. Catalysts, inorganic and organic (enzymes), stimulate this transfer of electrons. The rôle played by catalysts is not difficult to understand, if we accept this viewpoint, nor, in fact, are any other chemical reactions. In short, the electronation and de-electronation concept makes use of the modern theory of atomic structure in placing in a common category all chemical reactions, whether they be reactions of direct union, of displacement, or of electrolysis, or oxidation-reduction reactions involving complex organic substances. *The vehicle of change is the electron in all cases.*

CHAPTER XIII

DYNAMICAL PHYSICAL CHEMISTRY OF THE CELL

1. The Unit of Life is the Cell.—The bodies of animals and plants are made up of structural units called cells, which may be compared to the bricks of a building. Within these cells there is a substance of a distinctive nature, a substance that is found nowhere else. This substance is called *protoplasm* (*protos*, first; *plasma*, form). Surrounding each cell there is an elastic membrane termed the cell wall. In each organ we find a mass of these cells closely resembling one another in every respect. Each of these cells exhibits activities peculiar to living matter. These activities are manifested in the following ways:

(*a*) By movement, which takes the form of alternating contraction and expansion.

(*b*) By metabolism, i.e., by absorption of matter from the surrounding media; by the production of definite substances which accumulate in the protoplasm, but which undergo change and are finally exuded from the cell.

(*c*) By transmission of stimuli; i.e., chemical, electrical, and other kinds of stimuli received by one cell result in activities in a different cell.

(*d*) By reproduction, which takes the form of the division of the cell into two daughter cells.

It follows that the function of any organ is the sum of the activities of its constituent cells, and that the metabolism of a complex organism is the sum of the changes of the cells that compose it. It is therefore logical to study the functions and reactions which are common to all the structural units with a view to the application of the knowledge so gained to the elucidation of the more intricate problems of the body as a whole, in health and in disease.

From the data presented in the foregoing chapters, it is clear that all chemical reactions are modified by physical environment as well as by their chemical environment. Take, for example, the nature of the

container in which an enzyme is acting on a suitable substrate. The behavior of the enzyme in a glass container may be quite different from what it would be in a semi-permeable container submerged in a solution of electrolytes. So, in the living cell, whose structure permits of the diffusion of some substances while preventing that of others, and where the medium is exceedingly complex and variable, the action of an enzyme is not always the same in intensity, or even in type of reaction. The same enzyme will sometimes hydrolyze a complex molecule into simpler substances, and under other conditions will synthesize the complex molecule from the simpler constituents.

2. The Chemical and Physical Characteristics of the Cell.—We have seen that the cell is the structural unit of the living organism, and that within it all the changes that affect life must take place. The mass of these cells consists of protoplasm, which may be regarded as made up of an aggregation of colloids and crystalloids, electrolytes and non-electrolytes, dissolved in water and in each other.

A microscopical examination of cell protoplasm shows that it is not a simple homogeneous mass, but that it exhibits a definite structure, being composed of (a) a small dark portion termed the nucleus; (b) numerous granular bodies of various sizes and kind; and (c) a clear mass of more or less liquid material, which if examined with an ultramicroscope will be observed to possess the typical characteristics of a colloidal emulsoid or hydrogel. The whole mass, in practically all plant tissues, is surrounded by a clear-cut cell-wall; in animals, the membrane covering of the cell protoplasm is not generally as easily detected, yet there is positive proof of its universal existence.

Free animal cells (red blood corpuscles) and many plant cells, if placed in a solution of lower osmotic pressure that their contents, will tend to swell, owing to endosmosis; or, if they are placed in a hypertonic solution, they will usually shrink, because of loss of water. (For further data on this topic, review the chapter on osmosis and diffusion.)

The cell wall differs quite a little from the cell contents in chemical composition as well as in physical state. The difference, however, is quantitative rather than qualitative, especially in animal tissues. In general, it may be stated that certain substances are found in every living cell. These are termed the *primary* constituents of the cell. They include the following: proteins—albumins, globulins, nucleo-proteins, phospho-proteins, glyco-proteins; lipins—phospho-lipins, sterols, etc.; carbohydrates—glucose, pentose, and possibly glycogen;

electrolytes—the various salts of potassium, sodium, calcium, and magnesium. Many other *secondary* constituents may also be present. These include several proteins that are not soluble in water or dilute saline solutions; starch, glycogen, cellulose, and several of the less common mono-, di-, and tri-saccharids, glucosids, etc.; iron, manganese, iodine, etc.; and simple fats.

It is evident that the secondary constituents, such as cellulose or the insoluble proteins, will tend to accumulate in the cell membrane; but in many animal tissues the cell membrane very closely resembles the cytoplasm (the non-nuclear portion of the cell contents) in composition, the main difference being a lesser water content.

The cell membrane must not be considered as a separate container in which the protoplasm has been placed. It is an actual part of the living cell, having a similar metabolism to the cytoplasm and nucleus, and dying when the rest of the cell dies.

3. Permeability of Cell Membranes.—It has already been pointed out (p. 146) that living cells are surrounded by a membrane, or surface layer which acts as a membrane, which possesses the property of allowing the passage in and out of the solvent (water) at all times, and of some electrolytes and non-electrolytes at certain times, but prevents the passage of colloids. It has also been shown that cell membranes tend to prevent the equal distribution of certain ions on both sides of the membrane. In other words, cell membranes possess not only the property of *semi-permeability*, but also that of selective permeability. This property of selective permeability is usually lost after the cell dies. It is evident, then, that the living cell membrane only partially resembles the artificial semi-permeable membranes employed in determinations of osmotic pressure. Of course, artificial membranes can be made of practically any degree of permeability, but they do not possess the property of adaptability to their surroundings which is so characteristic of living tissue membranes. The latter are self-forming, variable, and adaptable to the ever-changing conditions of living matter, being quite permeable to a dissolved substance under one set of conditions and impermeable under other conditions.

That the protoplasmic substance of the cell is capable of forming a new membrane is shown by the fact that when pure protoplasm is brought in contact with a solution, through the rupture of the cell wall, it exhibits the same osmotic phenomena as the original cell. This was proved by Nageli in 1855 by the following experiment: A

root hair of *Hydrocharis* (a water plant with relatively long root hairs, which are processes of the root cells themselves) was placed under a cover-glass in a solution of dye, such as aniline blue, to which the normal cells are impermeable. The root hair was then crushed by pressure and, from the places where the cell wall was torn, masses of protoplasm exuded and formed little balls. These balls showed similar osmotic phenomena to those of the entire cell. The protoplasm remained unstained by the dye. Similar experiments have been performed and analogous results obtained by other investigators, notably Kuhne and Pfeffer.

4. Osmotic Pressure of the Tissue Fluids.—The diffusible constituents of living matter, and of the media that bathe it, play an important part in determining the movements and distribution of water in the living cells. In other words, osmotic pressure is one of the most important factors in controlling the distribution of water in protoplasm. It will be remembered that the concept known as osmosis is the passage of fluids, or dissolved substances, into or out of any tissue, or substance that is surrounded by a membrane of suitable composition. Since protoplasm is largely composed of hydrogels and hydrosols, the membrane controlling the osmosis may be either the cell wall or the internal colloidal films which are present in the entire mass of the cell contents.

In aqueous colloidal solutions, such as are found in living cells, there is a very definite surface tension manifested at the surface of the different interfaces of the suspended particles and the solvent. This has an important bearing on the phenomenon of "adsorption," which in turn plays a most important part in cellular dynamics.

The force which actuates the movement of water into or out of the tissue cells is the difference between the osmotic pressure of the cell contents and that of the external medium which bathes the cells. The manner in which this force may impel the migration of water will become clear if we make use of the osmotic phenomena discussed in previous chapters, especially Chapters VII, VIII, and XI. The peculiar factors in the osmotic pressure of tissue fluids are due to the fact that the membrane of the living cell is not only self-forming, but alters in its permeability during the life processes. It may at one time allow a given solute to pass through, and at another time prevent its passage. Another peculiarity of the cell membrane is that it seems endowed with the power of choosing which constituents of the sur-

rounding fluid may pass in, or which of its internal constituents may pass out. Sometimes comparatively large particles can enter, whereas smaller particles of other kinds are kept out. Moreover, the process of *negative osmosis* may take place through a cell membrane; that is, water may pass from the more concentrated side to the more dilute side under the peculiar activities of living cells. This flow of water takes place against osmotic pressure.

5. The Properties of Water and Cellular Dynamics.—Water is one of the absolutely essential constituents of protoplasm, for a variety of reasons. In the first place, it is the most efficient solvent for the various constituents of protoplasm, and we have evidence to show that all the energies which are displayed in the living cell are derived from the energies of substances in aqueous solution. Then, again, water is the most efficient ionizing solvent (with a couple of unimportant exceptions); hence, solution in water permits electrical forces to come into play. The high specific heat of water is also of value to living organisms, since the average cell contains about 80 per cent water, and when heat is liberated in metabolic reactions the temperature of the cell does not rise very greatly, the water playing the rôle of a heat regulator, that is, taking up the heat which is liberated and giving it off gradually. The high surface tension of water is essential in the conservation of the boundaries of the cell, and in their restoration after displacements due to cellular activities. Furthermore, in living tissues, practically all the chemical reactions that take place within the cells are reversible hydrolyses; i.e., water actually enters into the reaction or is liberated by it, and the equilibrium point of the reaction varies according to the proportions of water within the cells. Even the oxidation-reduction reactions involve the ions of water in the process, or water is involved in the exchange of electrons that takes place in the procedure. It is also interesting to note that the percentage of water is highest in those tissues that are undergoing the most rapid metabolic changes. Also, the percentage of water is lower in adult than in embryonic tissues, and decreases with advancing age of the tissues and slowing up of the speed of metabolism.

6. Surface Phenomena of Cells.—In the chapter on the colloidal state of matter, the phenomena of surface boundaries were presented. It was pointed out that when any substance exists in the colloidal state it has a relatively large surface area, and that, consequently, the phenomena of surface tension are exhibited at the interfaces between

the colloidal suspension and the solvent. Since protoplasm is known to exist largely in the form of a colloidal gel, it is more than probable that such surfaces exist in protoplasm between the more solid and the more liquid parts of the protoplasm; and that the physical properties of such surfaces of separation play an important part in the biophysical properties of cells. It is clear that the molecules that are in the surface layers at the various interfaces, where the two phases of a colloidal system are in contact, are not under similar attractions in all directions. The molecules in the surface layer are exposed on the inner side to the attraction and influence of one set of molecules, while those on the other side are exposed to the influence of a different aggregation of molecules. The result of this condition is that the outward force exerted upon the molecules in the surface film will be different from the inward force exerted upon them. Hence, there develops at the surface a state of tension which greatly affects the chemical reactivity of the molecules that are present in this surface layer, and so the surface layer comes to possess different chemical and physical properties from the interior, the most evident change being restriction in molecular movement, which results in the acquisition of the properties of a solid. All surfaces have a definite resistance to rupture, owing to the fact that the molecular freedom of movement is restricted in certain directions; and this resistance to rupture of the surface film is known as surface tension. For the discussion of methods of determining surface tension, see page 266 *et seq.*

The surface tension of a liquid is altered by dissolved substances. All solutes, with the exception of certain inorganic salts, lower the surface tension. It will be recalled (see p. 269) that Gibbs and Thomson showed that dissolved substances which lower surface tension tend to concentrate at the surface on account of the fact that the free energy of the system is thereby lessened. The adsorbed substance often forms a continuous layer of molecules; and, if this layer is of a solid nature, there will form a solid film on the surface. Sometimes a second layer may be added to the first, and in this way a relatively tough composite membrane may be built up.

Many surfaces are the seat of an electric charge. Most substances that are insoluble in water are negatively charged in relation to water. This charge affects the adsorption which takes place on the surface of such particles, and will tend to favor the concentration of electropositive as against electronegative ions at the interface. When a solute

has an electric charge, either as an ion or as a colloidal particle, and the surface in contact with the solution also has a charge, the degree of adsorption depends on the nature of the charge carried by the two substances; no decrease of free energy would be produced by adsorption of a negatively charged particle on a surface carrying the same charge, but the reverse. On the other hand, adsorption of an oppositely charged particle tends, through neutralization of the charge, to a diminution of free energy. Since electrolytes may augment or diminish the charge at a surface, it is evident that there must always be a proper adjustment of the constituents of solutions to the properties of the surface of the living cells with which they may come in contact.

7. Nature of the Cell Membrane.—We have seen that the living cell membrane differs from the artificial semi-permeable membrane in the possession of ability to change its permeability with the changes in its environment. These alterations in permeability may be due to a variety of causes, such as electrical, chemical, light, and heat stimuli; the stimulus of functional activity; and the addition of any adsorbable substance.

Before considering the factors that alter the permeability of cell membranes, it is desirable to summarize a few of the established facts concerning the nature of such membranes. In the first place, it seems certain that the membrane consists of a complex colloidal system of all the cell constituents, together with those of the outer medium, the most important character-giving constituents of animal cell membranes, however, being proteins and lipoids.

Secondly, the membrane must not be looked upon as an invariable permanent structure, for the forces that effect its formation will tend, when removed or reversed, to deform it. When a cell dies, the semi-permeability of its membrane is usually entirely lost.

Thirdly, the semi-permeability of living cell membranes is of necessity never absolute. This becomes obvious when we consider that the materials forming the cell membrane must be obtained from the constituents of the cell, which in turn are dependent upon the ever-changing nutrient materials dissolved in water. Unless a cell can be penetrated by the various substances out of which protoplasm is built up, the progressive consumption of material and dissipation of energy by the cell must soon lead to disintegration.

Fourthly. Nernst[1] has shown by experimentation that the per-

[1] Zeit. phys. Chem., 1891.

meability of a membrane for a given substance is essentially dependent upon the solvent power of the membrane for the said substance. This point, which is of the greatest importance in the study of permeability phenomena, has been thoroughly investigated by Overton.[2]   He studied the behavior of the living cells toward dyestuffs and many other substances which are insoluble or only slightly soluble in water, but readily soluble in fats or fat-like bodies. As a result of this invest- igation, Overton proposes his well-known "lipoid theory" of per- meability of cell membranes. Substances that are soluble in lipoids can pass through the living cell membrane, but substances that are insoluble in lipoids cannot.

According to Overton, the protoplasm-boundary layer is probably impregnated with *lipoids*, i.e., phospho-lipins (lecithin, cephalin, etc.), sterols (cholesterol, phytosterol, etc.), glyco-lipins (cerebron, phren- osin, etc.), sulpho-lipins, and amino-lipins. Lipoids are not a class of bodies that can be clearly defined chemically; they resemble the simple fats in regard to solubility and solvent power, but only some of them resemble them even partially from a chemical standpoint. Some of these bodies are always present in protoplasm and to a greater extent in the cell membrane. They help to give to protoplasm its ability to contain a large amount of water without dissolving; and also the power of taking up readily and in relatively large amounts many substances readily soluble in fats but not readily soluble in water.

For the proper understanding of the rôle of the phospho-lipins, the most essential lipoids from a physico-chemical standpoint, it is neces- sary that one should be familiar with their chemical structure. This topic is, of course, treated in considerable detail in the standard texts on biological chemistry, but for our purpose it may be sufficient to bear in mind that, in addition to possessing great structural similarity to the neutral tri-glycerides (simple fats), they also possess one free acidic group attached to a phosphoric acid radical, which in turn is attached to a basic nitrogen group. These substances, like the proteins, are amphoteric, and hence they have the power to unite with either acids or bases. They have an iso-electric point, at which they un- doubtedly show a greater physical instability than at other hydrogen- ion concentrations, and they can form salt-like combinations with acids at a pH below this point, and with bases above it. The fact that one rarely obtains a phospho-lipin free from small amounts of inorganic

[2] Vier. d. Natur. Ges. in Zurich, **44**, 1899.

matter (potassium, calcium, magnesium, etc.) suggests that such combinations are common in living cells.

There is a reasonable presumption that the living cell membrane is composed essentially of a protein-phospho-lipin complex. Such a layer, at all events, has been the basis of a great many investigations and speculations concerning cell permeability. It was formerly thought that the precipitating action of anions on phospho-lipins followed the famous Hofmeister series, originally observed in the case of proteins; but recent experiments with these lipins at definite pH concentrations show that they obey the simple valency rule in an analogous manner to the proteins (p. 290). We now have evidence to show that lipoids, especially the phospho-lipins, as well as proteins, when present in an aqueous solution containing electrolytes, such as are present in the living cell, will form a two-phase system, and, therefore, the ionic equilibria at the inter-surfaces of such heterogeneous systems are likely to be controlled by the Donnan membrane equilibrium theory (see p. 295).

Although considerable support for the "lipoid solubility" theory of cell membranes has been forthcoming, there is little doubt that a mere fatty pellicle, whether or not admixed with proteins and other substances, is not adequate to account for the variation in permeability that is characteristic of the living cell membrane. For one thing, it is very difficult to understand how the permeability can be regulated by the processes occurring inside and outside of the cell unless some other factors assist in the process. But, whatever explanation may be given, it seems certain that the cells show peculiar permeability to substances soluble in lipoids and lipoid solvents, whereas they are only at times permeable to those not so soluble, e.g., sugars, amino acids and most salts. There are, it would appear, two kinds of permeability, of which the latter alone is subject to functional change.

Ruhland,[3] after making a thorough investigation of the permeability of cell membranes to many substances, advanced strong evidence that, at least as regards dyes and enzymes, permeability is not a question of solubility in the membrane, but of the dimensions of the particles or molecules; that is, the membrane may be looked upon as a sieve. Similar ideas have been advanced by Bartell[4] and by

[3] Kolloid-chemische Protoplasmastudien, Koll. Zeit., 12, 113, 1913.
[4] J. Phys. Chem., 16, 318, 1912.

Duclaux and Errera.[5] These authors conceive the properties of semi-permeable membranes to depend on the size of the pores and the character of their walls. The principal property of the walls that they studied extensively was the electric charge. An electronegative membrane tends to repel anions, and an electropositive one to repel cations. In brief, these authors believe that the size of the pores plus the electrical properties of the cell walls are the factors that control the passage of large particles through cell membranes.

8. The Emulsion Structure of Protoplasm.—Clowes[6] has shown that there is a rather close resemblance between certain conditions found in living tissues and those observed in complex emulsions. We have already seen (p. 280) that there are two types of emulsions—emulsions of oil in water and emulsions of water in oil. It has been shown that sodium oleate emulsifies oil in water, while calcium oleate emulsifies water in oil; a mixture of the two oleates behaves differently, depending on the relative amounts. It has also been found that there is a definite ratio of calcium to sodium at which the two oleates will practically balance each other, and the slightest relative change will modify the type of emulsion. Clowes,[7] also working with a sodium soap and calcium chloride, obtained a similar result. He believes that there is a very close resemblance between the effects produced in such physical systems and those produced in biological systems. Interesting data along this line were obtained by Jacques Loeb, who showed that certain marine organisms died in sodium chloride or calcium chloride solutions which were isosmotic with sea water, but flourished when there was a definite ratio of sodium to calcium, a result which could not be explained on the basis of osmotic pressure. Clowes found that the ratio of sodium and calcium salts necessary to produce a balance between the two types of emulsions (mentioned above) when working with oil, water, and soap, was practically the same as that found in sea water. This indicates that the connection between the two sets of phenomena is remarkably close.

If we consider protoplasm as consisting essentially of lipoids and a watery protein solution containing various electrolytes, we shall have an emulsion of oil in water in the presence of an excess of sodium salts, and of water in oil in the presence of an excess of calcium salts.

[5] Rev. Gen. Colloids, **2**, 130, 1924.
[6] J. Biol. Chem., **24**, 1916.
[7] J. Phys. Chem., **29**, 407, 1916.

On the basis of these and analogous experiments, Clowes has advanced an interesting theory regarding cell membranes. He suggests that the cell membrane is a system consisting of a watery solution of proteins and lipoids. When the cell is in the resting state, the membrane consists of an emulsion of protein solution in lipoid, the latter being the dispersion medium. In this state, he claims, the membrane will be permeable only to lipoid-soluble substances; but, when the cell is in the active state, the reverse type of emulsion is obtained, i.e., the watery protein solution becomes the dispersion medium, and then water-soluble substances are able to pass through.

This view of Clowes, that the cell membrane is a complex emulsion system of colloidal substances, the two phases of which are capable of undergoing reversal, according to whether the cell is resting or active, offers a plausible explanation of many perplexing biological problems, and there are many data to back it up.

9. Alterations in Permeability of the Cell Membrane.—That the permeability of a cell may be altered under the influence of various stimuli, has been known for many years. Nageli, as far back as 1855, observed that when plant cells died chlorophyll and other substances diffused out of them. However, changes in permeability are also known to occur without the death of the cell. Such changes may be brought about by some external stimulus or by alterations in the material surrounding the cell.

It has been found that light increases the permeability of plant cells to most electrolytes, and, to a lesser degree, to non-electrolytes like sugar and urea.

X-rays and violet light rays affect the permeability of both plant and animal cells.

The composition of the boundary surface of cells depends largely on the substances present within the cell and in the surrounding media, plus the nature of the electrical charges of the interface. Any alteration in any of these factors will tend to alter the permeability. For example, Osterhout[8] found that the permeability of Spirogyra cells is increased by pure solutions of NaCl, but inhibited by the addition of $CaCl_2$; and Fluri[9] found that a three days' immersion in a 0.01 per cent solution of aluminium sulphate makes Spirogyra cells permeable to most salts as well as to glucose, and that the cells can be brought back

---

[8] Science, **34**, 187, 1911.
[9] Flora, **99**, 81, 1909.

to their normal state of permeability by placing them in pure water. Again, Siebeck[10] showed that frog's muscle, if immersed in isotonic potassium chloride, swells, thereby showing that the action of the potassium salt tends to diminish or abolish the permeability to the potassium ion, which the muscle normally possesses in the presence of mixed electrolytes. Gerard[11] found that, on feeding animals with excess of potassium salts, the blood maintains its normal composition, while the cells of the tissue lose sodium. These data prove that changes in ionic environment may bring about changes in the permeability of a cell membrane.

Adsorption is another factor which changes the permeability of the cell membrane. If a solute of opposite electrical sign to the membrane comes into contact with it, adsorption will take place, and the physicochemical nature of the membrane will be altered. The membrane may be thickened, the interstices may be wholly or partially occluded, or the electrical charge may be wholly or partially neutralized. It is not inconceivable that all of these changes may take place in the same membrane.

The permeability of the active cell differs considerably from that of the resting cell. The resting muscle cell, for example, is impermeable to many substances that readily pass in or out of the active cell. The resting muscle cell membrane is impermeable to glucose and amino acids; but during activity of the cell the membrane becomes permeable to these substances, and remains so as long as the tissue is functionally active. This change in permeability is undoubtedly due to some local stimulus, such as a change in the hydrogen-ion concentration; or, it may be that during the activity of the cell a specific ion is liberated which then acts upon the membrane, changing its permeability. A singular result might be obtained by the temporary removal of a given ion through chemical combination or adsorption. It can be shown experimentally, for example, that muscle cells are entirely altered in their permeability if placed for a time in a fluid free from Ca salts, and under these conditions may become permeable to substances such as KCl, which they do not ordinarily allow to pass. It is evident, then, that calcium ions are necessary in order to maintain the normal permeability of the cell membrane.

10. Electrical Changes in Protoplasm.—The material composing living tissues and cells is permeated throughout with electrolytes in

[10] Pflüger's Arch., **150**, 316, 1913.   [11] Compt. rend., **154**, 1912.

very dilute solution, hence in a high degree of dissociation. Further-more, the cell protoplasm contains in solution or suspension various colloids which are usually electrically charged. In fact, there is evidence to show that every activity of living matter is accompanied by an electrical disturbance in the protoplasm of the tissues involved.

Colloidal particles, whether in solution, in suspension, or in membranes, tend to adsorb electrolytes; but they do not adsorb all electrolytes with the same degree of intensity, hence the electrical state of the material on either side of a membrane is continually undergoing alterations. Every functional change in a tissue, therefore, tends to induce a difference in electrical potential in different parts of the tissue. Experimental proof of this electrical disturbance is easily obtained. If any two parts of a normal resting muscle are connected by a delicate galvanometer in a circuit, no current will be observed. If the living tissue, however, is undergoing different activities, e.g., if one part of a muscle is injured, a deflection of the galvanometer will be observed. A similar result will be obtained if the instrument is connected with a voluntary muscle that is undergoing irregular contractions. In short, it has been found that every heart beat, every contraction of a muscle, every chemical change, and every exhibition of glandular secretion is associated with a change in electrical potential.

In all the experiments that have been conducted along this line, the following two facts have been generally noted:

(1) The part of the tissue that is most active is always electronegative to the part that is less active.

(2) The current always flows from the more active to the less active portion of the tissue.

In making a model of the plasma membrane of the living cell, in order to study the E.M.F., McClendon[12] suggests that it is desirable to have the electrolyte responsible for the change of such a nature and concentration that it may exist in the living tissue. Several ions have been suggested, but McClendon favors H^+. To use his own words: "Hydrogen ions, being so very rapid in diffusion and so small that they might easily pass through membranes, seem very likely to be the ions concerned, especially as the outside of the cell is electropositive to the inside and acids (carbonic and lactic) are produced inside the cell because of the metabolism of the cell. The hydrogen-ion concentration

[12] Physical Chemistry of Vital Phenomena, Princeton, 1917.

on the outside of many cells in the body corresponds to a pH of about 7.45 (the blood plasma being 7.45). In order to produce an electromotive force of about 60 millivolts it would be necessary to have the hydrogen-ion concentration on the inside of the cell correspond to a pH of about 6.45. The acidity of this solution is then less than that of ordinary distilled water exposed to the CO_2 of the atmosphere. The pressed juice of many tissues is more acid than this. Therefore it seems possible that the difference in concentration of hydrogen ions inside and outside the cell is sufficient to give rise to an electromotive force of about 60 millivolts. This is approximately the electromotive force that is usually observed in the so-called current of injury by which the difference in potential on the cut and uncut surfaces, that is, inside and outside of the cell, is measured. Therefore, a model was set up containing a membrane of benzene. Buffer solutions were placed on the two sides of the membrane, one having a pH of about 7 to represent the outside of the cell, and the other of about 6 to represent the inside of the cell. An electromotive force of 31 millivolts was obtained."

The stimulated portion of a cell becomes electronegative, and this electronegativity is transmitted to other portions of the cell. Thus, almost instantly, the electropositivity of the cell surface is regained, the portion that was negative first regaining its positivity first. In some plant cells the time elements are longer and may be several minutes (McClendon).

11. The Reaction of Protoplasm and the Preservation of Neutrality.—In several of the preceding chapters, data concerning the acidity and alkalinity of biological media and tissues were presented, and in Chapter X several up-to-date methods for the determination of these reactions were cited; hence, in the present chapter, only data concerning the living cell will be presented.

Owing to the difficulties involved in working with living cells, indirect methods are usually employed in the determination of the reaction. It is quite common, for instance, to determine the reaction of the fluid in which the cells live. It is not likely that a cell will have a reaction very different from the medium in which it lives. At any rate this method gives results that are sufficiently accurate for many purposes.

Perhaps the most remarkable fact that has been discovered about protoplasm and protoplasmic fluids is the constancy of their reaction, which remains almost neutral, being usually just slightly alkaline,

despite the fact that a large proportion of the materials that are brought in contact with them are either acid or alkaline in reaction. The maximum range of variation of reaction in the blood of normal individuals, for instance, is indicated by pH 7.30–7.50, yet the blood is continually receiving large quantities of acids and alkalies. (For the physico-chemical explanation of the above phenomena, review Chapter IX.)

It is a significant fact, however, that the slight changes that do take place in the reaction of protoplasm have very profound effects. The first effect induced by a change in the reaction toward acidity in protoplasm is a reduction in the rate of oxidation. The respiratory oxidation of the cell is very much affected by slight changes in the reaction. A decrease in alkalinity slows up the rate of respiration, while an increase in alkalinity augments it. Following a decrease in oxidation caused by decreased alkalinity, it is always found that a number of suboxidation products tend to accumulate in the system. Thus, in *diabetes* the faulty oxidation of fats produces a quantity of non-volatile acids which cannot be discharged, as carbon dioxide is discharged, through the respiratory epithelium of the lungs. These bodies may disappear almost entirely shortly after the alkalinity is brought back to normalcy.

A very important effect of change in the reaction of protoplasm is the alteration of the enzymotic activity of the cell. In the last chapter is was pointed out that the activity of enzymes is greatly dependent on the hydrogen-ion concentration of the medium. An increase in pH concentration augments the activity of some enzymes and retards that of others; it activates many dormant enzymes and reverses the action of others. Acidity stimulates the carbohydrate-splitting enzymes, which increases the supply of easily oxidizable carbohydrates, thereby tending to enable the cell to recover its metabolic balance. Again, an increase in acidity stimulates the activity of the proteolytic enzymes, thereby effecting an increase in ammonia by augmenting the hydrolysis of the proteins of the cell. The increase in acidity is thus checked almost as soon as it develops, owing to the great sensitivity of the biological catalysts to changes in the reaction of their environment.

Another physical change produced by a slight increase of hydrogen ions in the cells is the imbibition of water by the colloidal material and the resultant swelling of the tissue. In diseases that cause acidosis,

this becomes a very serious complication. (For a discussion of the general concept of imbibition, see pages 307–311.)

Finally, it will be recalled that most of the constituents of protoplasm are "buffers," i.e., substances that tend to hold the (H^+) and (OH^-) concentrations of a solution at a constant value. In living tissues and tissue fluids, proteins, amino acids, carbonates, and phosphates all function as buffers, and therefore help to regulate the reaction.

AUTHOR INDEX

SUBJECT INDEX